

GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2004, 13:13:40 ; Search time 3339 seconds

(without alignments)

6786.981 Million cell updates/sec

Title: US-09-669-817A-4

Perfect score: 523

Sequence: atctgtatccatggatcacca.....gaccgcaagtggaaatggaa 523

Scoring table: IDENTITY_NUC

GapOp 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1: gb_baa:*

2: gb_htg:*

3: gb_inn:*

4: gb_omt:*

5: gb_ovr:*

6: gb_Pat:*

7: gb_phi:*

8: gb_Pti:*

9: gb_Pri:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_uni:*

14: gb_vir:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_on:*

21: em_or:*

22: em_ovr:*

23: em_Pat:*

24: em_ph:*

25: em_pl:*

26: em_xo:*

27: em_sts:*

28: em_un:*

29: em_vir:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rid:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

RESULT 1

AKU03829

LOCUS AKU03829

DEFINITION Orzya sativa (Japonica cultivar-group) cDNA clone:033148A06, full

inset sequence.

ACCESSION AKU03829

VERSION AKU03829.1 GI:32899038

KEYWORDS FLI_CDNA; CAP traper.

ORGANISM Orzya sativa (Japonica cultivar-group)

Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrartoideae; Oryzeae.

REFERENCE 1

The Rice Full-Length cDNA Consortium, National Institute of

Pred. No. is the number of results predicted by chance to have a

Agrobiochemical Sciences Rice Full-Length cDNA Project Team:
 Kikuchi,S., Sato,K., Nagai,T., Kawagashira,N., Doi,K.,
 Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
 Koima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
 Onitsuka,K., Shinkai,T., Foundation of Advancement of International
 Science Genome Sequencing & Analysis Group: Otomo,Y., Murakami,K.,
 Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tunoda,Y., Xie,Q., Lu,M.,
 Kurosoki,T., Kodama,T., Maudia,H., Kobayashi,M., Xie,Q., Lu,M.,
 Nariyama,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
 Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
 Kusumegi,T., Okamoto,Y., Ryu,R., Ueda,M., Matsumura,K., RIKEN,
 Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
 Hara,A., Hasidume,W., Hayatsu,N., Imamura,K., Ishii,Y., Ichinomori,
 Kasawa,I., Kondo,S., Konno,H., Miyazaki,A., Otao,N., Ota,Y.,
 Saito,R., Sasaki,D., Sato,K., Shibusawa,K., Shiragawa,A., Shiraki,T.,
 Yoshino,M. and Hayashizaki,Y.
 Collection, mapping, and annotation of over 28,000 cDNA clones from
 Japonica rice
Science 301 (5631), 376-379 (2003)
 MEDLINE
 22752273
 PUBLMED
 12869764
 2 (bases 1 to 1819)
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
 Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
 Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoaka,T.,
 Horii,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
 Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
 Kawagawa,S., Katoh,H., Kawagashira,N., Kawai,M.,
 Kikuchi,S., Kishiwada-Hirokawa, Kishimoto,N., Kobayashi,M.,
 Kadama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kuida,M.,
 Maeda,S., Kurihara,C., Kuroda,T., Kusunose,T., Li,C., Du,M.,
 Mizuno,K., Matsubara,K., Matsuya,T., Miura,J., Miyazaki,A.,
 Mizuno,K., Murakami,K., Murata,M., Nagao,T., Nakamura,M.,
 Namiki,T., Nariyama,R., Niikura,J., Nishi,K., Nomura,K.,
 Numasaki,R., Ohneda,M., Ohsukai,K., Okamoto,Y., Oka,H.,
 Osto,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
 Sakazumi,N., Sano,H., Sasaki,D., Sato,K., Shibata,K.,
 Shinagawa,A., Shiroki,T., Shishiki,T., Socabe,Y., Sugano,S.,
 Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
 Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
 Toya,T., Tsunoda,Y., Ueda,M., Waki,Q., Yahagi,W.,
 Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
 Yoshihara,A.
 Direct Submission
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
 Agrobiological Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
 Tel: +81-29-838-7007, Fax: +81-29-838-7007)
 This clone is one of the 28K full-length cDNA clones from Japonica
 rice.
 URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
 NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Sato,K.,
 Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
 Ishikawa,M., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
 Oneda,E., Ohneda,M., Suzuki,K., Li,C., Onitsuka,K., Shishiki,T.,
 Yamamoto,M.,
 FASIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
 Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
 Kodama,T., Kurosoki,T., Kusumegi,T., Ito,M., Matsumura,K., Miura,J.,
 Mizuno,K., Narikawa,R., Niikura,J., Otao,N., Ueda,M., Xie,Q., Yokomizo,S.,
 Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
 Yoshimura,A., Matsubara,K. and Murakami,K.,
 Genome Exploration Research Group in Riken Genomic Sciences Center
 and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
 Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
 Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
 Hiraoaka,T., Horii,F., Iida,J., Immamura,K., Imamura,K., Ishii,Y.,
 Itoh,M., Kagawa,I., Karagawa,S., Katoh,H., Kawai,J., Kuida,M.,
 Kishikawa-Hirokawa,T., Kojima,Y., Konno,S., Kouda,M.,
 Koya,S., Kuribayashi,C., Matsuura,M., Miyazaki,A., Murata,M.,
 Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Otao,N.,
 Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazumi,N., Sano,H.,
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 PUBLISHED
 JOURNAL
 SCIENCE 301 (5631), 376-379 (2003)

FEATURES

Source /organism="Oryza sativa (japonica cultivar-group)" /mol-type="mRNA" /cultivar="Nipponbare" /db xref="taxon:33947" /clone="U033148A06"

Location.Qualifiers

1. 1819

Query Match 98.5%; score 515; DB 8; length 1819;
Best Local Similarity 99.0%; Pred. No. 1.8e-66; Mismatches 5; Indels 0; Gaps 0;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

Db 1 ATCTGATCAGCGATACCAAGTCTCTCACCTCCACTCCAGGGCAGCTGACAGA 60
17 ATCAGATCAGCAGATACCAAGTCTCTCACCTCCACTCCAGGGCAGCTGACAGA 60

QY

Db 61 GAGCTCCACCGCGGCCAACATGTTCTCACCCACCGTACGCCACCTCTGAGC 120
77 GAGCTCCACCGCGGCCAACATGTTCTCACCCACCGTACGCCACCTCTGAGC 136

QY

Db 121 ACCGCTTCACCTCCATCTTGCCCTGAGCTCCCTTCTGTGACCTTCCCG 180
137 ACCGCTTCACCTCCATCTTGCCCTGAGCTCCCTTCTGTGACCTTCCCG 196

QY

Db 181 CGCTTCTCCACCTCCATCTTGCCCTGAGCTCCCTTCTGTGACCTTCCCG 240
197 CGCTTCTCCACCTCCATCTTGCCCTGAGCTCCCTTCTGTGACCTTCCCG 256

QY

Db 241 ATCTCACCCCTCCACCAAGGCCAACCTAGCTCCCTCTGACCCCTTCTCC 300
257 ATCTCACCCCTCCACCAAGGCCAACCTAGCTCCCTCTGACCCCTTCTCC 316

QY

Db 301 ACCCCCCTCACCCACCGCTCTCCAACTCGAGCTCCCTCGCCGCCACACC 360
317 AGACCTCACCGACCGGTCTCCACCTCGAGCTCCCTCGCCGCCACACC 376

QY

Db 361 CGCGGCCACCGCAAGAGCACCATTGACCCAGTCACCGCCGCTGGTAAGT 420
377 CGAGGCCACCGCAAGAGCACCATTGACCCAGTCACCGCCGCTGGTAAGT 436

QY

Db 421 GGACCAAGGAGAACCTGAGGGGGAGAGAGTGCTCAAGGGGGAGCTGAGCTG 480
437 GGACCAAGGAGAACCTGAGGGGGAGAGAGTGCTCAAGGGGGAGCTGAGCTG 496

QY

Db 481 ACTCCCCCTACGAGCATGGATCGACCGCAAGTGGAGTGGGA 523
497 ACTCCCCCTACGAGCATGGATCGACCGCAAGTGGAGTGGGA 539

RESULT 2

AP001129/_C

AP001129 Locus AP001129 Definition 194509 bp DNA linear PLN 21-MAR-2002
Accession AP001129 Definition ORYZA sativa (japonica cultivar-group) genomic DNA, chromosome 6,
Organism Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Keywords Sporophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Reference 1 Sasaki, T., Matsumoto, T. and Yamamoto, K., Shinagawa, A., Shiraki, T., Takahashi, F.,
Authors Sogabe, Y., Tegami, M., Tagami-Takeda, Y., Tagawa, A., Takachi, A., Tomaru, A., Toya, T., Waki, K.,
Title Yasunishi, A. and Hayashizaki, Y.
Journal Published Only in Database (2000)
Reference 2 (bases 1 to 194509)

Ov	481	ACTCCCCCTACGACGATGATTGACCGCAGTGGAAAGTGGGA	523
/translation	"MAQSLALDGSVPPSRIVSERQAGLPRRFMPESATGRETVTLG		
/transl_start	"MAQSLALDGSVPPSRIVSERQAGLPRRFMPESATGRETVTLG		
/transl_end	"MAQSLALDGSVPPSRIVSERQAGLPRRFMPESATGRETVTLG		
Db	17165	ACTCCCTACGACGATGATTGACCGCAGTGGAAAGTGGGA	17123
RESULT	3		
DEFINITION	Orzya sativa (japonica cultivar-group) cDNA clone:001-027-D03, full		
ACCESSION	AK060638		
LOCUS	AK060638		
VERSION	AK060638.1		
KEYWORDS	GI:32970656 LI_CDNB; oligo capping.		
ORGANISM	Orzya sativa (Japonica cultivar-group)		
REFERENCE	1		
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team;		
TITLE	Kikuchi,N., Sato,K., Nagata,T., Kawagashira,N., Ooka,H., Hotta,I., Kishimoto,N., Yasaki,J., Igihara,M., Yamada,H., Oka,H., Hotta,I., (22/2208)"		
JOURNAL	Kojima,K., Namiki,T., Ohneda,E., Yangi,W., Suzuki,K., Li,C.,		
MEDLINE	Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science, Genome Sequencing & Analysis Group., Ohno,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,		
PUBLMED	Kurusaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Ito,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Isahaki,J., Kawamura,M., Yoshimura,A., Mura,J., Kusunegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;, Kawai,J., Carninci,P., Adachi,J., Aizawa,T., Arakawa,T., Fukuda,S., Hara,A., Hashidome,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Tomaru,A., Ota,Y., Shiraki,T., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Yoshino,M., and Hayashizaki,Y.		
REMARKS	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
COMMENT	This clone is one of the 25 full-length cDNA clones from japonica rice.		
URL	http://cdna01.dna.affrc.go.jp/CDNA/		
Query Match	Best Local Similarity 98.5%; Score 515; DB 8; Length 194509;		
Matches	518; Conservative 0; Mismatches 5; Indels 0; Gaps 0;		
Qy	1 ATCTGTACGAGATGCCAAGTCCTCTCACTCCACTTGACGGAGCTGAGACA	60	
Db	17645 ATCAGATCAGAGATACCAAGTCCTCTAACCTCCACTTGACGGAGCTGAGACA	17586	
Qy	61 GGCTCCACCCGGCCGGCCAGAGTCCTCACCAACGGTACCCACCTCTCGAGC	120	
Db	17585 GAGCTCCACCCGGCCGGCCAGATGAGTCACCCACCGTAGCCACCTCTCGAGC	17526	
Qy	121 ACCCTCTCCCTCCAAGCGCTCTCTCCCTCGCCCTCTCCACCG	180	
Db	17525 ACCCTCTCCCTCCAAGCGCTCTCTCCCTCGCCCTCTCCACCG	17466	
Qy	181 CCTTCTCCACCTGCCATCTCCCTGACTCCCTCGTGTGACCTTCC	240	
Db	17465 CGCTCTCCACCTGCCATCTCCCTGACTCCCTCGTGTGACCTTCC	17406	
Qy	241 ATCTCCACCCCTCTCCACCGCCGACCATAGCTCCCTCGACCCCTTCCTCC	300	
Db	17405 ATCTCCACCCCTCTCCACCGCCGACCATAGCTCCCTCGACCCCTTCCTCC	17346	
Qy	301 ACACCTCACCAGGACCTCTCCACCTGCCGACCCACCC	360	
Db	17345 ACACCTCACCAGGACCTCTCCACCTGCCGACCCACCC	17286	
Qy	361 CCAGGCCACCTGCCGACCTCTCCACCTGCCGACCCACCC	420	
Db	17285 CCAGGCCACCTGCCGACCTCTCCACCTGCCGACCCACCC	17226	
Qy	421 GACCCACGAGCAACGCTCTCCACCTGCCGACCCACCC	480	
Db	17225 GACCCACGAGCAACGCTCTCCACCTGCCGACCCACCC	17166	

	NTAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawashiria,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,I., Hotta,I., Kojima,K., Namiki,T., Oneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., and Yamamoto,M.	LOCUS	AK120505
	FAIR Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamura,M., Kobayashi,M., Koda,M., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narasaki,R., Niimura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K., and Murakami,K.	DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:j013123a20, full insert sequence.
	Genome Exploration Research Group in Riken: Adachi,J., Aizawa,K., and Genome Science Laboratory in Riken: Fukuda,P., Carninci,P., Hara,A., Hashizume,W., Hayashida,K., Hiraoka,T., Horii,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itch,M., Kawagishi,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koyama,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ora,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Tagami,M., Tagami,Y., Tagami,Y., Tagawa,A., Takahashi,F., Yasukawa-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A., and Hayashizaki,Y.	DEFINITION	Oryza sativa (japonica cultivar-group)
FEATURES		ACCESSION	AK120505_1 GI:37990128
source		VERSION	FLI cDNA; CAP traper.
		KWIDWORD	Oryza sativa (japonica cultivar-group)
		ORGANISM	Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
ORIGIN		REFERENCE	1
		The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagishi,M., Yamada,H., Oka,H., Hotta,I., Kojima,K., Namiki,T., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurobaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Nikura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsuura,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kawawa,J., Kondo,S., Konno,H., Matsuura,M., Osaoto,N., Orai,I., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M., and Hayashizaki,Y.	
		TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from Japonica rice
		JOURNAL	Science 301 (5631), 376-379 (2003)
		MEDLINE	12869764
		PUBMED	
		REFERENCE	2
		AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanasaki,T., Hara,A., Hashizume,W., Hayashida,K., Hotta,I., Iida,J., Iida,Y., Ichii,Y., Ichikawa,M., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katchi,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuya,T., Mura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niimura,J., Nishi,K., Nomura,K., Numasaki,R., Oneda,B., Oinou,M., Otsuka,K., Oka,M., Ooka,H., Osoyo,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Samo,H., Sasaki,D., Sato,K., Sato,K., Sato,K., Sato,K., Sato,K., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,T., Sugiyama,Y., Tagami,T., Tagami-Y., Tagami-Y., Tagawa,A., Takahashi,F., Yasukawa-Akahira,S., Xie,Q., Yano,Y., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yoshimura,A., and Yoshimura,A.
		JOURNAL	Unpublished
		REFERENCE	3 (bases)
		AUTHORS	Kikuchi,S.
		TITLE	Direct Submission
		JOURNAL	Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamionada, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@affrc.go.jp)
		COMMENT	Tel: 81-29-838-7007, Fax: 81-29-838-7007) This clone is one of the 32K full-length cDNA clones from japonica rice. URL: http://cdna01.nia.affrc.go.jp/cDNA/
RESULT	NTAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawashiria,N., Doi,K., Kishimoto,N., Yazaki,J., Namiki,T., AK120505/C		

	insert sequence.
ACCESSION	AK063929
VERSION	AK063929.1
KEYWORDS	GL:297947
ORGANISM	Oryza sativa (japonica cultivar-group)
REFERENCE	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Koijima,S., Namiki,T., Ohneda,E., Yanagi,W., Suzuki,K., Iii,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kuroasaki,T., Kodama,T., Matsuda,H., Kobayashi,M., Xie,Q., Lu,M., Nariyama,R., Suyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kubumugi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Ishii,Y., Itoh,M., Hara,A., Hashizume,W., Hayatsu,N., Itoh,M., Kurogawa,T., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M., and Hayashizaki,Y.
AUTHORS	Eukaryota; Viriliplanta; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzeae; Oryza.
1	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Koijima,S., Namiki,T., Ohneda,E., Yanagi,W., Suzuki,K., Iii,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kuroasaki,T., Kodama,T., Matsuda,H., Kobayashi,M., Xie,Q., Lu,M., Nariyama,R., Suyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kubumugi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Ishii,Y., Itoh,M., Hara,A., Hashidome,W., Hayatsu,N., Itoh,M., Kurogawa,T., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M., and Hayashizaki,Y.
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL	Science 301 (5631), 376-379 (2003)
MEDLINE	22752273
PUBMED	12869764
REFERENCE	(bases 1 to 1548)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Hara,A., Hashidome,W., Hayatsu,N., Hizamoto,K., Hirao,T., Hotta,I., Iida,Y., Iida,Y., Kawamata,M., Itoh,M., Kegawa,I., Horii,F., Hotta,I., Iida,Y., Iida,Y., Kawagashira,N., Kawai,J., Kawamata,M., Imotani,K., Ishibiki,J., Iishi,Y., Ishikawa,M., Kikuchi,S., Kishikawa-Hirazane,T., Kishimoto,N., Kobayashi,M., Koijima,S., Koijima,T., Kojo,M., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kuroasaki,T., Kubumugi,T., Li,C., Iii,M., Matsuda,H., Matsubara,K., Matsuya,M., Miura,J., Miyazaki,A., Mizuno,K., Murakawa,R., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Nariyama,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Oheda,E., Ohno,M., Ohtsuki,K., Oka,M., Oka,H., Oshiro,N., Ota,Y., Otomo,Y., Ryu,R., Saito,H., Sakai,C., Sakai,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshihara,A.
TITLE	Submitted (05-DCC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice.
URL	http://cdna01.dna.affrc.go.jp/cDNA/
PIGS	NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Imai,A., Oka,H., Hotta,I., Koijima,T., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Iii,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FUJIMURA	PIGS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kuroasaki,T., Kurokawa,R., Kubumugi,T., Lu,M., Matsuda,H., Miura,J., Mizuno,K., Nariyama,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,

FEATURES		TITLE JOURNAL	
source	Query Match Best Local Similarity 58.9%; Matches 269; Conservative 0; Db	Direct Submission Submitted (15-AUG-2001) Takuji Sasaki, National Institute of Biological Sciences, Rice Genome Research Program; Kannondai 2-12, Tsukuba, Ibaraki 305-8602, Japan (E-mail: tasaki@nias.afric.go.jp, URL: http://rgp.dna.afric.go.jp/), Takao-Akahira, S., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Yasunishi, A. and Hayashizaki, Y.	
LOCATION/QUALIFIERS		/organism="Oryza sativa (japonica cultivar-group)" /mol-type="mRNA" /clone="J033115E12";	
1. 1565			
FEATURES	SOURCE	COMMENT	On Nov 26, 2003 this sequence version replaced gi:1508417. The orientation of the sequence is from Mirke to -21M13 of the BAC clone. The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data.
ORIGIN		/organism="Oryza sativa (japonica cultivar-group)" /mol-type="genomic DNA" /cultivar="Nipponbare" /db-xref=taxon:39947" /chromosome="2" /clone="OJ1212_C06"	
RESULT 7	AP004049	Query Match Best Local Similarity 58.9%; Matches 269; Conservative 0; Db	ORIGIN
LOCUS	AP004049	Score 130.2; DB 8; Length 1565;	Query Match Best Local Similarity 58.9%; Matches 269; Conservative 0; Db
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,	Score 130.2; DB 8; Length 110989;	Query Match Best Local Similarity 58.9%; Matches 269; Conservative 0; Db
ACCESSION	AP004049	PLN 27-NOV-2003	Score 130.2; DB 8; Length 110989;
VERSION	AP004049.2	linear	PLN 27-NOV-2003
KEYWORDS	GI:38664198	DNA	linear
SOURCE	Oryza sativa (Japonica cultivar-group)	chromosome 2,	DNA
ORGANISM	Oryza sativa (Japonica cultivar-group)		chromosome 2,
HTG	BAC clone:OJ1212_C06, complete sequence.		
RESULT 8	AP004851	Query Match Best Local Similarity 58.9%; Matches 269; Conservative 0; Db	ORIGIN
LOCUS	AP004851	Score 13189 bp.; DB 8; Length 110989;	Query Match Best Local Similarity 58.9%; Matches 269; Conservative 0; Db
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,	Score 13189 bp.; DB 8; Length 110989;	Query Match Best Local Similarity 58.9%; Matches 269; Conservative 0; Db
ACCESSION	AP004851	HTG 23-MAR-2002	Score 13189 bp.; DB 8; Length 110989;
VERSION	AP004851.1	linear	HTG 23-MAR-2002
KEYWORDS	HTG; Rfns PHASE2.	DNA	linear
SOURCE	Oryza sativa (Japonica cultivar-group)		DNA
ORGANISM	Oryza sativa (Japonica cultivar-group)		
HTG	Eukaryota; Viriplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
	Birhartoideae; Oryzeae; Oryza.		
REFERENCE	1		
AUTHORS	Sasaki, T.; Matsunoto, T. and Yamamoto, K.		
TITLE	Oryza sativa nippobare (Ga3) genomic DNA, chromosome 2, BAC		
JOURNAL	Published Only in Database (2001)		
REFERENCE	2 (bases 1 to 110989)		
AUTHORS	Sasaki, T.; Matsunoto, T. and Yamamoto, K.		
TITLE	Oryza sativa nippobare(Ga3) genomic DNA, chromosome 2, BAC		

JOURNAL	clone:OJ1359_D06
PUBLISHED ONLY IN DATABASE	(2002)
2 (bases 1 to 135189)	
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	Submitted (20-MAR-2002) Takiji Sasaki, National Institute of Agricultural Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail: tsasaki@niasaffrc.go.jp, URL: http://rgp.dna.affrc.go.jp/ , Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT	The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data.
NOTE:	It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
* NOTE:	This is a 'working draft' sequence.
* THIS SEQUENCE WILL BE REPLACED	*
* BY THE FINISHED SEQUENCE AS SOON AS IT IS AVAILABLE AND	*
* THE ACCESSION NUMBER WILL BE PRESERVED.	
FEATURES	
SOURCE	
ORIGIN	
Query Match	24.9%; Score 130.2; DB 2; Length 135189;
Best Local Similarity	58.9%; Pred. No. 1.8e-10;
Matches	269; Conservative 0; Mismatches 173; Indels 15; Gaps 2;
QY	79 CCACCATGAGTCCTACCCACCGTAAAGCCACCTCTCACGACCCCTCTCCATC 138
Db	28963 CCACCTTCCAACTCTCTCTCTCTCTCTCTCTCTCCACACCCACATTCTCC 2022
QY	139 CACGGCGGRCCT 198
Db	29023 CT 20082
QY	199 CATTCTTCCCCTGACTCCCCCTTGCTGCTGACCCCTTCATCTCCACCCCTTCC 258
Db	29083 GCTTCACT 29142
QY	259 CCAG-----CGACCGTAGTCCCTCTCACCTTCTCTCTCACCC 306
Db	29143 ACCTTGACCTCTCTCTCCAACTCCGACCCACCTGTAACTCTCTCTCTCT 29202
QY	307 TCACGGACGGCGCTCTCCAACTCGAGCTGGCTGGCGCCACACCCAGGC 366
Db	29203 GCGATCTCTCTCTCTCTCTCCACACTCTCTCTCTCTCTCTCTCTCTCT 29262
QY	367 CCACAGCGAGGTGACCTATGTACGGAGTACCGGCCGATGGTCAGTGACCA 426
Db	29263 CCCACGCGACAATACACTACGAGGCGAACGAGATCAAGGACCA 29322
QY	427 CGAGGACAGCTGGCTGGCGGGAGGGTGCTCAAGTGGAGGTGAATGAGGCA 486
Db	29323 CCCAGACAGACAGCC---CGAAGGGCGGACCTTAAGTGGAGGGCCAGATCGACACC 29379
QY	487 CCTGGACATGGATCTGACGGCAAGTGAATGAGTGA 523
Db	29380 CCAAGGACAGATGGCTTCGAGCGCAATGTGGAGTGA 29416
RESULT	10
ACCESSION	AC022663/c
LOCUS	AC022663
DEFINITION	Homo sapiens clone RP11-2905, LOW-PASS SEQUENCE SAMPLING.
VERSION	AC022663
KEYWORDS	AC02263.2 GI:913958 HTG; HTGS PHASE0.
SOURCE	Homo sapiens (human)
ORGANISM	Bukrysta, Metacea; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Buteraria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 63082)
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE	Homo sapiens, clone RP11-2905
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 63082)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckley,R., Beda,F., Castle,A., Boguslavskiy,L., Bouknight,B., Brown,A., Burkett,G., Castle,A., Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Coote,P., DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenster,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardina,S., Grant,G., Haiges,B., Heaford,A., Horton,L., Karatas,A., Klein,J., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
RESULT	9
DEFINITION	Sequence 14 from patent US 5670367. DNA linear PAT 28-DBC-1997
LOCUS	166494
VERSION	166494.1
KEYWORD	GI:2724471
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 7218)
AUTHORS	Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.
TITLE	Recombinant foamy virus
JOURNAL	Patent: US 5670367-A 14-23-SEP-1997;
FEATURES	source 1 . 7218 /location/Qualifiers
source	/organism="unknown"
/mol_type="unassigned DNA"	
ORIGIN	
Query Match	20.9%; Score 109.7; DB 6; Length 7218;
Best Local Similarity	3.2%; Pred. No. 4.2e-07;
Matches	12; Conservative 260; Mismatches 98; Indels 0; Gaps 0;
QY	23 TCTCTTCACATCCACTCCACGGCACTGAGAGAGCTACCCGGCCAC 82
Db	1085 YY 1144
QY	83 CATCAGTGTCAACCCGGTACGCCACCTCTCTGACGAGACCCCTCTCCAC 142
Db	1145 YYY 1204
QY	143 GCCTCTCTCCCTCTGCTGCCTCTCTCTCTCTCCACCCGCGTTCCTCCTC 202
Db	1205 YYY 1264
QY	203 CTTCGCCCCGACTCCCCCTTGCTGCTGACCCCTTCATCCTCACCCCTTCTCCAC 262
Db	1265 YYY 1324
QY	263 GCGACGACTAGCTCCCTCTGACCCCTCTCTCACACCCCTACGGGGTC 322
Db	1325 YYY 1384
QY	323 CAACTGAGCTCCCTCTGCGACCCGCGCCACACCCAGGCCACCCAGGAGTG 382
Db	1385 YYYGTTACCAATT 1444
QY	383 CACCTATGTC 392
Db	1445 CTTCTATCTC 1454
RESULT	10
ACCESSION	AC022663
LOCUS	AC022663
DEFINITION	Homo sapiens clone RP11-2905, LOW-PASS SEQUENCE SAMPLING.
VERSION	AC022663
KEYWORDS	AC02263.2 GI:913958
SOURCE	HTG; HTGS PHASE0.
ORGANISM	Bukrysta, Metacea; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Buteraria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 63082)
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE	Homo sapiens, clone RP11-2905
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 63082)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckley,R., Beda,F., Castle,A., Boguslavskiy,L., Bouknight,B., Brown,A., Burkett,G., Castle,A., Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Coote,P., DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenster,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardina,S., Grant,G., Haiges,B., Heaford,A., Horton,L., Karatas,A., Klein,J., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,

JOURNAL
 Landers, T., Lehmkuhl, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
 Macdonald, P., Marquis, N., McElhaney, P., McGurk, A., McKernan, K.,
 McPheeeters, R., O'Farrell, T., O'Donnell, P., Morrow, J., Nayler, J.,
 Norman, C.H., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
 Roy, A., Santos, R., Severy, P., Spender, B., Strange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talmas, J., Tesfaye, S., Theodore, J.,
 Tizell, A., Vassiliev, R., Vie, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
 Zimmer, A. and Zody, M.

TITLE
 Direct Submission

COMMENT
 Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6910643.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1995-1997)
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wim.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L4854
 Center clone name: 29_0_5

* NOTE: This record contains 77 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 734: contig of 734 bp in length
 * 735 834: gap of 100 bp
 * 835 1556: contig of 722 bp in length
 * 1557 1656: gap of 100 bp
 * 1657 2352: contig of 696 bp in length
 * 2452 3148: contig of 696 bp in length
 * 3149 3248: gap of 100 bp
 * 3249 3945: contig of 697 bp in length
 * 3946 4045: gap of 100 bp
 * 4046 4728: contig of 683 bp in length
 * 4729 4828: gap of 100 bp
 * 4829 554: contig of 720 bp in length
 * 5549 5648: gap of 100 bp
 * 5649 6362: contig of 714 bp in length
 * 6363 6462: gap of 100 bp
 * 6463 7171: contig of 709 bp in length
 * 7172 7271: contig of 100 bp
 * 7272 7997: contig of 726 bp in length
 * 7998 8097: gap of 100 bp
 * 8098 8835: contig of 738 bp in length
 * 8836 8935: gap of 100 bp
 * 8936 9664: contig of 729 bp in length
 * 9665 9764: gap of 100 bp
 * 9765 1015: contig of 751 bp in length
 * 10615: gap of 100 bp
 * 10616 1142: contig of 814 bp in length
 * 11529: gap of 100 bp
 * 11530 12294: contig of 765 bp in length
 * 12295 12395: gap of 100 bp
 * 12395 13177: contig of 782 bp in length
 * 13277 13991: contig of 715 bp in length
 * 13992 14091: gap of 100 bp
 * 14092 14497: contig of 706 bp in length
 * 14498 14798: gap of 100 bp
 * 14898 15607: contig of 710 bp in length

15608 15707: gap of 100 bp
 15708 16432: contig of 725 bp in length
 16433 16532: gap of 100 bp
 16533 17269: contig of 736 bp in length
 17269 17368: gap of 100 bp
 17369 18105: contig of 737 bp in length
 18106 18205: gap of 100 bp
 18206 18888: contig of 683 bp in length
 18889 18988: gap of 100 bp
 18988 19702: contig of 714 bp in length
 19703 19802: gap of 100 bp
 19803 20516: contig of 714 bp in length
 20517 20616: gap of 100 bp
 20617 21330: contig of 714 bp in length
 21331 21430: gap of 100 bp
 21431 22143: contig of 713 bp in length
 22144 22243: gap of 100 bp
 22244 22950: contig of 707 bp in length
 22951 23050: gap of 100 bp
 23051 23785: contig of 735 bp in length
 23786 24605: contig of 721 bp in length
 24607 24707: gap of 100 bp
 24707 25447: contig of 741 bp in length
 25448 25548: gap of 100 bp
 25548 26276: contig of 729 bp in length
 26276 26377: gap of 100 bp
 26377 27107: contig of 731 bp in length
 27108 27208: gap of 100 bp
 27208 27927: contig of 719 bp in length
 27927 28027: gap of 100 bp
 28027 28738: contig of 711 bp in length
 28738 28837: gap of 100 bp
 28838 29559: contig of 700 bp in length
 29560 29659: gap of 100 bp
 29659 29660: contig of 715 bp in length
 29660 30374: gap of 100 bp
 30375 3074: contig of 710 bp in length
 3074: gap of 100 bp
 30475 31184: contig of 710 bp in length
 31185 31284: gap of 100 bp
 31284: gap of 100 bp
 31285 32011: contig of 727 bp in length
 32012 3211: gap of 100 bp
 3211: gap of 100 bp
 32112 32836: contig of 725 bp in length
 32836: gap of 100 bp
 32336: gap of 100 bp
 32336: contig of 685 bp in length
 32337 32937: gap of 100 bp
 32937 33721: contig of 727 bp in length
 33721 3444: contig of 723 bp in length
 3444: gap of 100 bp
 34445 34545: contig of 709 bp in length
 34545 35254: gap of 100 bp
 35254 35353: gap of 100 bp
 35353 36669: contig of 716 bp in length
 36669: gap of 100 bp
 36669 36778: contig of 709 bp in length
 36778: gap of 100 bp
 36778 36879: contig of 709 bp in length
 36879 36978: gap of 100 bp
 36978 37677: contig of 699 bp in length
 37677: gap of 100 bp
 37677 37777: gap of 100 bp
 37777 37778: contig of 712 bp in length
 37778 38489: gap of 100 bp
 38489: gap of 100 bp
 38589: contig of 708 bp in length
 38589 39597: contig of 708 bp in length
 39597: gap of 100 bp
 39597 39397: contig of 708 bp in length
 39397: gap of 100 bp
 39397 39298: contig of 728 bp in length
 39298 39398: contig of 728 bp in length
 39398 40125: gap of 100 bp
 40125: gap of 100 bp
 40226 40225: contig of 720 bp in length
 40225 40345: contig of 720 bp in length
 40345: gap of 100 bp
 40445 41045: gap of 100 bp
 41045 41767: contig of 722 bp in length
 41767: gap of 100 bp
 41767 41868: contig of 736 bp in length
 41868 42003: contig of 718 bp in length
 42003: gap of 100 bp
 42003 42604: contig of 717 bp in length
 42604 42704: gap of 100 bp
 42704 4320: contig of 703 bp in length
 4320: gap of 100 bp
 4320 43521: contig of 703 bp in length
 43521: gap of 100 bp
 43521 44223: contig of 703 bp in length
 44223: gap of 100 bp
 44223 44324: contig of 718 bp in length
 44324: gap of 100 bp
 45042 45141: gap of 100 bp

* 45142 43855: contig of 714 bp in length
 * 45956: gap of 100 bp
 * 45955: contig of 724 bp in length
 * 46680 46779: gap of 100 bp
 * 47482: contig of 703 bp in length
 * 47483 47582: gap of 100 bp
 * 48318 48317: contig of 735 bp in length
 * 48418 49147: contig of 730 bp in length
 * 49148 49247: gap of 100 bp
 * 49248 49267: contig of 720 bp in length
 * 49968 50067: gap of 100 bp
 * 50068 50772: contig of 705 bp in length
 * 50873 51568: contig of 696 bp in length
 * 51568 51668: gap of 100 bp
 * 51669 52399: contig of 731 bp in length
 * 52400 52499: gap of 100 bp
 * 52500 53224: contig of 725 bp in length
 * 53225 5324: gap of 100 bp
 * 53325 54013: contig of 689 bp in length
 * 54014 54113: gap of 100 bp
 * 54114 54344: contig of 731 bp in length
 * 54845 5444: gap of 100 bp
 * 55660 55660: contig of 716 bp in length
 * 55661 55760: gap of 100 bp
 * 55761 55690: contig of 730 bp in length
 * 56491 56591: gap of 100 bp
 * 56591 57333: contig of 743 bp in length
 * 57334 57333: gap of 100 bp
 * 58161 58161: contig of 728 bp in length
 * 58162 58261: gap of 100 bp

Query Match
 Best Local Similarity 20.7%; Score 108.; DB 2; Length 63082;
Matches 221; **Conservative** 0; **Mismatches** 243; **Indels** 0; **Gaps** 0;

<pre> 15 CACCAAGTCCTCTTCAACTCCACTTCAGGGAGCTGAGAGAGCTAACCGC 75 36038 CCCTCACCCCCCCCCCCCCCCCCACCCCCCACCCCCCCCCCCCCCCCC 35979 76 CGGCCACCATGAGTTTACACCGTTACGCCACCTCTTGACGCCCTCTCTCC 135 35978 CCCCTCC 35919 136 TCCCAAGCT 195 35918 CCCCTCC 35859 b 196 ACCATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 255 35858 CCCCTCC 35799 256 TCCCAAGGCCGACACTAGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 315 35798 CCCCTCC 35739 316 GCGTCCTCCAACTCTGAGCTGCCTCTGGCGGCCACCCACCCAGGCCA 375 35738 CCCCTCC 35679 376 GGAAGTCACCTATGTCACCGAGTCCACCGGCCGATGGTCAAGTGGACC 435 35618 CCNNNNNNCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 35619 436 AGCTCGCGCGCGAGAGGGCTCAAGTAGGAGAGGTGACTC 479 35618 ANCCTCACCTGGGCCCTCAGGTCTCTGTC 35575 </pre>	<p> DEFINITION Mus musculus chromosome 16 clone RP23-197M9, WORKING DRAFT SEQUENCE, 101 unordered pieces. ACCESSION AC084799 VERSION AC084799.1 GI:11192127 KEYWORDS HUG, HGVS PHASE1, HGVS_DRAFT SOURCE Mus musculus (house mouse) ORGANISM Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; REFERENCE 1 (bases 1 to 303091) JOURNAL Unpublished AUTHORS Doe, Joint Genome Institute. TITLE Sequencing of Human Chromosome 16 COMMENT Submitted (17-NOV-2000) Production sequencing Facility, DOB Joint Center Project Name: 0 Center Clone name: RPCI-23-197M9 REFERENCE -----Genome Center, Joint Genome Institute Center Code: JGI Web site: http://www.jgi.doe.gov ----- Project Information Center Project Name: 0 Center Clone name: RPCI-23-197M9 Summary Statistics Consensus quality: 152568 bases at least Q40 Consensus quality: 15579 bases at least Q30 Consensus quality: 189549 bases at least Q20 Estimated insert size: 198300; agarose-fp estimation Estimated insert size: 293051; sum-of-contigs estimation Quality coverage: 4.85 in Q20 bases; agarose-fp estimation Quality coverage: 3.28 in Q20 bases; sum-of-contigs estimation * NOTE: This is a 'working draft' sequence. It currently consists of 101 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. <table border="0"> <tr> <td style="vertical-align: top;"> <pre> 1 1116: contig of 1116 bp in length * 1117 1216: gap of unknown length * 1217 2236: contig of 1320 bp in length * 2537 2836: gap of unknown length * 2637 4057: contig of 1421 bp in length * 4058 4157: gap of unknown length * 4158 5245: contig of 1088 bp in length * 5246 5345: gap of unknown length * 5346 6461: contig of 1116 bp in length * 6462 6551: gap of unknown length * 6562 7725: contig of 1164 bp in length * 7726 7943: contig of 1318 bp in length * 7826 9343: gap of unknown length * 9144 9244 10492: contig of 1249 bp in length * 10492 11493: gap of unknown length * 11493 11666: contig of 1074 bp in length * 11666 11767 12050: contig of 1184 bp in length * 11767 12050: gap of unknown length * 12050 12950: gap of unknown length * 12950 13051 14285: contig of 1239 bp in length * 13051 14285: gap of unknown length * 14285 14386 15578: contig of 1093 bp in length * 14386 15578: gap of unknown length * 15578 15579 16779: contig of 1201 bp in length * 15579 16779: gap of unknown length * 16779 16878: contig of 1399 bp in length * 16878 18279 18378: gap of unknown length * 18279 18378: gap of 1074 bp in length * 18378 18437 19452: contig of unknown length * 19452 19452: gap of unknown length </pre> </td> </tr> </table> </p>	<pre> 1 1116: contig of 1116 bp in length * 1117 1216: gap of unknown length * 1217 2236: contig of 1320 bp in length * 2537 2836: gap of unknown length * 2637 4057: contig of 1421 bp in length * 4058 4157: gap of unknown length * 4158 5245: contig of 1088 bp in length * 5246 5345: gap of unknown length * 5346 6461: contig of 1116 bp in length * 6462 6551: gap of unknown length * 6562 7725: contig of 1164 bp in length * 7726 7943: contig of 1318 bp in length * 7826 9343: gap of unknown length * 9144 9244 10492: contig of 1249 bp in length * 10492 11493: gap of unknown length * 11493 11666: contig of 1074 bp in length * 11666 11767 12050: contig of 1184 bp in length * 11767 12050: gap of unknown length * 12050 12950: gap of unknown length * 12950 13051 14285: contig of 1239 bp in length * 13051 14285: gap of unknown length * 14285 14386 15578: contig of 1093 bp in length * 14386 15578: gap of unknown length * 15578 15579 16779: contig of 1201 bp in length * 15579 16779: gap of unknown length * 16779 16878: contig of 1399 bp in length * 16878 18279 18378: gap of unknown length * 18279 18378: gap of 1074 bp in length * 18378 18437 19452: contig of unknown length * 19452 19452: gap of unknown length </pre>
<pre> 1 1116: contig of 1116 bp in length * 1117 1216: gap of unknown length * 1217 2236: contig of 1320 bp in length * 2537 2836: gap of unknown length * 2637 4057: contig of 1421 bp in length * 4058 4157: gap of unknown length * 4158 5245: contig of 1088 bp in length * 5246 5345: gap of unknown length * 5346 6461: contig of 1116 bp in length * 6462 6551: gap of unknown length * 6562 7725: contig of 1164 bp in length * 7726 7943: contig of 1318 bp in length * 7826 9343: gap of unknown length * 9144 9244 10492: contig of 1249 bp in length * 10492 11493: gap of unknown length * 11493 11666: contig of 1074 bp in length * 11666 11767 12050: contig of 1184 bp in length * 11767 12050: gap of unknown length * 12050 12950: gap of unknown length * 12950 13051 14285: contig of 1239 bp in length * 13051 14285: gap of unknown length * 14285 14386 15578: contig of 1093 bp in length * 14386 15578: gap of unknown length * 15578 15579 16779: contig of 1201 bp in length * 15579 16779: gap of unknown length * 16779 16878: contig of 1399 bp in length * 16878 18279 18378: gap of unknown length * 18279 18378: gap of 1074 bp in length * 18378 18437 19452: contig of unknown length * 19452 19452: gap of unknown length </pre>		

19553	21012: contig of 1460 bp in length
21013	21112: gap of unknown length
21113	21169: contig of 1057 bp in length
22170	22269: gap of unknown length
22270	23367: contig of 1098 bp in length
23368	23467: gap of unknown length
23468	24559: contig of 1092 bp in length
24559	24659: gap of unknown length
24660	24825: contig of 1166 bp in length
25826	25925: gap of unknown length
25926	27036: contig of 1111 bp in length
27037	27136: gap of unknown length
27137	28245: contig of 1109 bp in length
28246	28346: gap of unknown length
28346	29056: contig of 1311 bp in length
29657	29755: gap of unknown length
29755	30833: contig of 1127 bp in length
30884	30883: gap of unknown length
30984	32125: contig of 1142 bp in length
32126	32225: gap of unknown length
32226	33197: contig of 1172 bp in length
33398	33497: gap of unknown length
33497	34565: contig of 1148 bp in length
34646	34645: gap of unknown length
34746	35595: contig of 1149 bp in length
35994	35994: gap of unknown length
35995	37156: contig of 1162 bp in length
37157	37257: gap of unknown length
38396	38396: contig of 1140 bp in length
38397	38497: gap of unknown length
38497	39571: contig of 1074 bp in length
39571	39571: gap of unknown length
39671	40852: contig of 1182 bp in length
40853	40952: gap of unknown length
40953	42116: contig of 1164 bp in length
42117	42216: gap of unknown length
42217	43317: contig of 1156 bp in length
43317	43472: gap of unknown length
43473	44103: contig of 1131 bp in length
44604	44703: gap of unknown length
44704	44954: contig of 1221 bp in length
44925	46024: gap of unknown length
46025	47174: contig of 1150 bp in length
47174	47274: gap of unknown length
47275	48744: contig of 1470 bp in length
48844	49989: gap of unknown length
49845	49989: contig of 1125 bp in length
49970	50059: gap of unknown length
50070	51135: contig of 1066 bp in length
51136	51235: gap of unknown length
51236	52630: contig of 1455 bp in length
52631	52691: gap of unknown length
52691	53920: contig of 1130 bp in length
53921	54020: gap of unknown length
54021	55101: contig of 1081 bp in length
55102	55201: gap of unknown length
55202	55324: contig of 1123 bp in length
55324	56324: gap of unknown length
56325	56425: gap of unknown length
56425	57583: contig of 1155 bp in length
57584	57683: gap of unknown length
57684	58832: contig of 1139 bp in length
58923	58932: gap of unknown length
58932	60547: contig of 1025 bp in length
60547	60649: gap of unknown length
60648	61743: contig of 1096 bp in length
61743	61843: gap of unknown length
61844	62998: contig of 1155 bp in length
62998	63098: gap of unknown length
63099	64220: contig of 1122 bp in length
64221	64320: gap of unknown length
64321	65398: contig of 1078 bp in length
65398	65498: gap of unknown length
65499	66762: contig of 1264 bp in length

```

Query Match 20.5%; Score 107; DB 0; Length 303091;
Best Local Similarity 51.9%; Pred. No. 3.7e-07;
Matches 182; Conservative 0; Mismatches 169; Indels 0;

```

Gaps

RESULT 14	QY	452 GAGGGTCTCAAGTGGAGGCTGAGCTCGACTCCCCCTACGACGATGGATT	503
LOCUS	AC079433/c	AC079433	252689 bp DNA linear HTG 01-SEP-2000
DEFINITION	Mus musculus chromosome 16 clone RP23-76F18, WORKING DRAFT	Mus musculus chromosome 16 clone RP23-76F18, WORKING DRAFT	SEQUENCE, 68 unordered pieces.
ACCESSION	AC079433	AC079433	AC079433.1 GI:9558045
VERSION		HGSC PHASEI; HGS, DRAFT.	HGSC PHASEI; HGS, DRAFT.
KEYWORDS		Mus musculus (house mouse)	
SOURCE			
ORGANISM		Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE		1 (bases 1 to 25689) 2 (bases 1 to 25689)	
AUTHORS		DOE Joint Genome Institute.	
TITLE		Sequencing of Mouse	
JOURNAL		Unpublished	
COMMENT		2 (bases 1 to 25689) DOE Joint Genome Institute.	
AUTHORS		Center Code: JGI Web site: http://www.jgi.doe.gov	
TITLE		Project Information	
JOURNAL		Center Project Name: 0 Center clone name: RPCI-23_76F18	
-----		Summary Statistics	
		Consensus quality: 136458 bases at least Q40	
		Consensus quality: 165595 bases at least Q30	
		Consensus quality: 177476 bases at least Q20	
		Estimated insert size: 18100; agarose-fp estimation	
		Estimated insert size: 245999; sum-of-contigs estimation	
		Quality coverage: 3.62 in Q20 bases; agarose-fp estimation	
		Quality coverage: 2.67 in Q20 bases; sum-of-contigs estimation	
		* NOTE: This is a 'working draft' sequence. It currently	
		* consists of 68 contigs. The true order of the pieces	
		* is not known and their order in this sequence record is	
		* arbitrary. Gaps between the contigs are represented as	
		* runs of N, but the exact sizes of the gaps are unknown.	
		* This record will be updated with the finished sequence	
		* as soon as it is available and the accession number will	
		* be preserved.	
		1 1137: contig of 1137 bp in length	
		1138 1337: gap of unknown length	
		2277: contig of 1040 bp in length	
		2278 2378: gap of unknown length	
		3185: contig of 1008 bp in length	
		3186: gap of unknown length	
		3486 4500: contig of 1075 bp in length	
		4560: gap of unknown length	
		4561 4661: contig of 1015 bp in length	
		5676 5775: gap of unknown length	
		5776 6905: contig of 1190 bp in length	
		6966 7055: gap of unknown length	
		7056 8274: contig of 1203 bp in length	
		8374: gap of unknown length	
		8375 9444: contig of 1090 bp in length	
		9454: gap of unknown length	
		9565 10799: contig of 1225 bp in length	
		10890 12400: contig of 1511 bp in length	
		12401 12500: gap of unknown length	
		12501: contig of 1300 bp in length	

Ruitz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C.,
 Shooshtari, N., Sisson, J., Sodergren, B., Sonalkar, T., Sparks, A.,
 Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerska, A.,
 Tamerisa, K., Tang, H., Tansey, J., Taylor, T., Telmissany, B.,
 Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D.,
 Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R.,
 Washington, C., Watlington, S., Williams, G., Williamson, A., Wu, Y.F., Zhou, J.,
 Wlezyk, R., Woerner, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J.,
 Zorrilla, S., Kucherlapati, R., Weinstock, G., and Gibbs, R.

JOURNAL Unpublished
COMMENT Direct Submission
REFERENCE 3 (bases 1 to 3281)
AUTHORS Worley, K.C.
TITLE Worley, K.C.
JOURNAL Direct Submission

Submitted (09-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: JGSI
 Center clone name: CH250-26N12

----- Summary Statistics
 Chemistry: Dye-primer Bodipy: int% of reads
 Chemistry: Dye-terminator Big Dye: int% of reads
 Consensus quality: 2 bases at least Q40
 Consensus quality: 26 bases at least Q30
 Consensus quality: 328 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/genbank_draft_data.html)
 * NOTE: The contigs are based on the application
 * of the PGI method using the Human genome (NCBI build 31)
 * as the comparative genome.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * by the accession number will be preserved.
 1. 3281: contig of 3281 bp in length.

FEATURES

source

/organism="Macaca mulatta"
 /mol_type="genomic DNA"
 /db_xref="taxon:9514"
 /clone="CH250-267N12"
 1. 3281
 /note="assembly name:CH250-267N12.1B
 /confidence: 0.67"

ORIGIN

Query Match 20.2%; Score 105.6; DB 2; Length 3281;
 Best Local Similarity 56.0%; Pred. No. 1.6e-06;
 Matches 200; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

15 TCACCAAGTCCTTACACTCCACTCCAGGGAGCTGAGCAGAGCTCACCGCG 74
 1776 TCCCTTGCGCCCTCCACCTCCGCCCGCCCGCCCTCCACCCCGCCCGCC 1717
 75 CGGGCAGCATGAGTTTACACCGTAACTCCACCTCTGAGAACCTTCTCCC 134
 1716 CCCCCCCCCCCCCCTCCACCTCCACCTCCACCCCTCCACCCCGCCCGCC 1657
 135 TTCCACCGCGCTCTCTCTCTCTGACCCCTCTCTCCACCGCGTTCCTCACC 194
 1656 CCCCCCCCCCCCCCCCCCTCCACCTCCACCCCGCCCGCCCGCCCGCC 1597

----- Search completed: April 14, 2004, 15:23:04
 Job time : 3348 secs

Copyright GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2004, 14:12:43 ; Search time 432 Seconds
(without alignments)
5143.073 Million cell updates/sec

Title: US-09-669-817a-4
Perfect score: 523
Sequence: 1 aactgtatcgatcgatccacca.....gaccgaaagtggaa 523

Scoring table: IDENTITY NUC
GapOp 10.0 , GapExt 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29jan04:
 1: geneseqn1980s:
 2: geneseqn1990s:
 3: geneseqn2000s:
 4: geneseqn2001as:
 5: geneseqn2001bs:
 6: geneseqn2002s:
 7: geneseqn2003as:
 8: geneseqn2003bs:
 9: geneseqn2003cs:
 10: geneseqn2004s:
 .

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
C 1	102	19.5	12733	6 ABK98631	Abk98631 Vector PE
C 2	102	19.5	12733	8 ACD1882	Acd1382 L. lactis
C 3	102	19.5	12739	6 ABK9593	Abk9592 Vector PE
C 4	102	19.5	12739	8 ACD1343	Acd1343 Plasmid P
C 5	101.4	3163	9 ADC07060	Adc07060 Human GPC	
C 6	98.2	18.8	5452	9 ADC08736	Adc08735 Human GPC
C 7	98	18.7	320	3 AAA38186	Aaa38186 Primer us
C 8	97.4	18.6	320	3 AAA38185	Aaa38185 Primer us
C 9	97.2	18.6	2188	2 AAZ77505	Aaz77505 Human ova
C 10	95.4	18.2	318	3 AAA38184	Aaa38184 Primer us
C 11	95.4	18.2	320	3 AAA38183	Aaa38183 Primer us
C 12	94.6	18.1	600	6 ABQ052497	Abq052497 Oligonucle
C 13	94.6	18.1	600	6 ABQ05496	Abq05496 Oligonucle
C 14	93.5	17.9	1065	6 ABT09882	Abt09682 Human Pau
C 15	92.2	17.6	1064	6 ABT09878	Abt09678 Human Pau
C 16	91.8	17.6	3133	9 ADC86738	Ad86738 Human GPC
C 17	89.4	17.1	9 ADC86598	Ad86598 Human GPC	
C 18	87.8	1337	2 AAZ1263	Aaz1263 Human gen	
C 19	86.8	1416	7 ABZ20567	Abz20567 Animal te	
C 20	85.8	14286	6 ABK9656	Abk9656 Mouse lisc	
C 21	84.8	1652	1000	3 AAB02484	Aab02484 Human col
C 22	83.8	1327	6 ABQ68452	Abq68452 Listeria	
C 23	81.8	840	6 ABQ35494	Abq35494 Oligonucle	

ALIGNMENTS

Result No.	Score	Query	Match Length	DB ID	Description
C 1	81.8	15.6	840	6 ABQ35495	Abq35495 Oligonucle
C 2	81.6	10.94	7 ACC78864	Acc78864 Oncolytic	
C 3	81.2	15.5	9 ADC86708	Adc86708 Human GPC	
C 4	80.8	15.4	4 AAI184343	Aai18433 Human pol	
C 5	80.8	15.4	4 AII47670	Abx47670 Bovine ES	
C 6	80.2	15.3	3 AAH18107	Aah18107 N. mening	
C 7	80	15.0	1050	4 AAD21685	Aad21685 Mutationa
C 8	78.6	15.0	4 AAD21684	Aad21684 Human ret	
C 9	78	14.9	6 AAD30228	Aad30228 Human PKD	
C 10	78	14.9	2 AAT194101	Aat194101 Human PKD	
C 11	78	14.9	2 AAT18551	Aat18551 Human pol	
C 12	78	14.9	2 AAT94108	Aat94108 Human PKD	
C 13	77.4	14.8	6 615	7 AAT23975	Aca23975 Prokaryot
C 14	77.4	14.8	6 615	6 ABP09669	Abt09669 Human PAL
C 15	76.8	14.7	3 AAT02504	Aaa0504 Human col	
C 16	76.6	14.6	3 AAI14663	Aai14663 Nucleotid	
C 17	76.6	14.6	3 AAI14661	Aai14661 Nucleotid	
C 18	76.6	14.6	3 AAI14664	Aai14664 Nucleotid	
C 19	76.6	14.6	3 AAI14662	Aai14662 Nucleotid	
C 20	76.6	14.6	3 AAI14651	Aai14651 Nucleotid	
C 21	76.4	14.6	2 AAB076213	Aab076213 HSV LysT	
C 22	76.2	14.6	6 ABQ70117	Abq70117 Listeria	

The invention relates to an isolated nucleic acid comprising a fusion promoter comprising at least one promoter that is modified to have altered activity in at least one gram-positive organism, or comprising altered activity in at least one gram-negative organism, or comprising PT in bacteria.

PT of xylo, teto, trp, malo or lambdaC, where at least one operator is

CC transcription so binding of a repressor to an operator represses
CC host cells comprising the fusion promoters, a method of identifying genes
CC involved in cellular proliferation or required for proliferation of a
CC prokaryotic cell using the vector, a method of identifying compounds that
CC inhibit the proliferation of a prokaryotic cell using the vector, a
CC method of identifying a compound that reduces the activity or level of a
CC gene product required for proliferation of a cell using the vector, a
CC compound identified by the methods, a method of inhibiting the activity
CC or expression of a gene in an operon required for proliferation using the
CC vector, manufacturing an antibiotic comprising using the vector or cell
CC and identifying a nucleic acid with promoter activity in *Enterococcus*
CC *faecalis*. The fusion promoters are useful for regulating nucleic acid or
CC polypeptide expression, particularly for regulating gene expression in
CC bacteria and for identifying proliferation-required genes or molecules
CC with potential antibiotic activity. The modified promoters are also
CC useful for replacing endogenous promoters to create cells with specific
CC regulatable genes. The present sequence is vector (or fragment)
CC incorporating a fusion promoter sequence of the invention

Example 3: Page 86-92: 1420c: English.
One plasmid that is believed to have antisense activity in at least one gram-positive organism, useful for regulating gene expression in bacteria.

Document has been received by the appropriate authority.

The invention relates to an isolated nucleic acid comprising a fusion promoter having at least one promoter that is modified to have altered activity in at least one gram-positive organism (e.g., *Staphylococcus aureus* or *Enterococcus faecalis*). The promoter is linked to at least one operator selected from xyLO, tetO, trpO, malO and lambda-cIOP, which are positioned such that the binding of at least one repressor to the operator represses transcription from the fusion promoter. Also included are a vector comprising the isolated nucleic acid, a host cell comprising the nucleic acid. The fusion promoter is useful for identifying genes involved in cellular proliferation, identifying a compound that reduces the activity or level of a gene product required for proliferation of a cell, inhibiting the activity or expression of a gene in an operon required for proliferation, manufacturing an antibiotic, identifying a gene that is required for proliferation of a prokaryotic cell, identifying a compound that inhibits the proliferation of a prokaryotic cell and regulating gene expression in bacteria. The present sequence is a plasmid containing a fusion promoter of the invention.

Db	4997	CCCCCC	4992	Query Match	Score	DB	Length
RESULT	3			Best Local Similarity	19.5%	6	12739
ID	ABK98592/c			Matches	54.9%	10;	
XX	ABK98592 standard; DNA; 12739 BP.			Pred. No.	3	8e-10;	
AC				Mismatches	0;		
XX	ABK98592;			Indels	165;		
DT	07-AUG-2003 (revised)			Gaps	0;		
DT	21-OCT-2002 (first entry)						
XX							
DE	vector PEPPERI containing XylR/Xylo/CP25 sequences.						
XX							
KW	ds; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;						
PS	P59; PLP; XYLO; teto; trpo; malo; lambda-c10; cellular proliferation;						
KW	antibiotic; vector.						
XX							
OS	Bubacteriia.						
OS	Bacteriophage lambda.						
OS	Escherichia coli.						
OS	Synthetic.						
XX							
PN	WO200251982-A2.						
XX							
PD	04-JUL-2002.						
XX							
PF	21-DEC-2001; 2001WO-US050250.						
XX							
PR	27-DEC-2000; 2000US-0259434P.						
PR	06-SEP-2001; 2001US-00948993.						
XX							
PA	(ELIT-) ELUTRA PHARM INC.						
XX							
PT	Haselbeck R, Wall D, Gross M;						
DR							
XX							
PS	WPI; 2002-575374/61.						
XX							
PT	isolated nucleic acid comprises bacterial promoters modified to have altered activity in at least one gram-positive organism, e.g. Bacillus anthracis or Clostridium botulinum, useful for regulating gene expression in bacteria.						
PT							
XX							
PS	Example 1; Page 206-209; 246pp; English.						
XX							
PT	The invention relates to an isolated nucleic acid comprising a fusion promoter comprising at least one promoter that is modified to have altered activity in at least one gram-positive organism, or comprising at least one operator consisting of Xy10, teto, trpo, malo or Lambda-c10, where at least one operator is positioned so binding of a repressor to an operator represses transcription from the fusion promoter. Also included are vectors and host cells comprising the fusion promoters, a method of identifying genes involved in cellular proliferation or required for proliferation of a prokaryotic cell using the vector, a method of identifying compounds that inhibit the proliferation of a prokaryotic cell using the vector, a method of identifying a compound that reduces the activity or level of a gene product required for proliferation of a cell using the vector, a compound identified by the methods, a method of inhibiting the activity or expression of a gene in an operon required for proliferation using the vector, manufacturing an antibiotic comprising using the vector or cell and identifying a nucleic acid with promoter activity in Enterococcus faecalis. The fusion promoters are useful for regulating nucleic acid or peptide expression, particularly for regulating gene expression in bacteria and for identifying proliferation-required genes or molecules with potential antibiotic activity. The modified promoters are also useful for replacing endogenous promoters to create cells with specific regulatable genes. The present sequence is vector (or fragment) incorporating a fusion promoter sequence of the invention. (Updated on 07-AUG-2003 to correct OS field.)						
SQ	Sequence 12739 BP; 3543 A; 2572 C; 3211 G; 3413 T; 0 U; 0 Other;						
PS	Example 1; Page 63-68; 142pp; English.						
RESULT	4			Query Match	19.5%	Score	102;
ID	ACD13843/c			DB	6;	Length	12739;
XX	ACD13843 standard; DNA; 12739 BP.			Best Local Similarity	54.9%;	Pred. No.	3
AC				Matches	201;	8e-10;	
XX				Mismatches	0;		
DT	15-AUG-2003 (first entry)			Indels	165;		
XX				Gaps	0;		
DE	Plasmid PEPPERI containing a Xy15-CP25 fusion promoter.						
XX							
KW	Promoter; ds; gram positive bacteria; staphylococcus aureus; Plasmid; Enterococcus faecalis; operator; xylo; teto; trpo; malo; lambda-c10; cellular proliferation.						
XX							
OS	Lactococcus lactis.						
OS	Synthetic.						
XX							
US	US2003027286-A1.						
XX							
PD	06-FEB-2003.						
XX							
PF	21-DEC-2001; 2001US-00032393.						
XX							
PR	06-SEP-2000; 2000US-0230335P.						
PR	27-DEC-2000; 2000US-0259434P.						
XX							
PA	(HASE/) HASELBECK R,						
PA	(WILL/) WALL D,						
PA	(GROS/) GROSS M.						
PT	Haselbeck R, Wall D, Gross M;						
DR							
XX							
PT	New isolated nucleic acid comprising a fusion promoter having at least one promoter that is modified to have altered activity in at least one gram-positive organism, useful for regulating gene expression in bacteria.						
PT							
SQ	Example 1; Page 63-68; 142pp; English.						

Db 137 ccc 78
 Qy 289 CCTCTCCTCACACCCCTCACGAGGGCTTCACACTGAGCTCCCTGAGGCC 348
 Db 77 ccc 18
 Qy 349 gcccccaaccccc 362
 Db 17 cccccccccccc 4

RESULT 8
 AAA38185/C
 ID AAA38185 standard; DNA; 320 BP.

Db 185 TCTCCCTACCTGCCATTCTGCCCTCACTCCACCTCTGACCCCTTCACT 244
 Qy 187 cc 128
 Db 245 CACACCCCTCTCCAGGCCGACCATGAGCTTCCTCCAGACCCCTTCTCCAC 304
 Qy 305 CTCTACCAGGCGCTCCAACTGAGCTCACTCACAGGAGGACACACCCAG 364
 Db 67 cc 8

Db DT 15-SEP-2003 (revised)
 DT 01-SEP-2000 (first entry)

DE Primer used in the analysis of a BVDV genome fragment.

XX Primer; bovine viral diarrhea virus; BVDV; nucleic acid analysis;
 KW diagnosis; pathological organism; detect; ss.
 XX OS Pestivirus type 1.

XX PN WO200020628-A1.

PD 13-APR-2000.
 PF 01-OCT-1999; 99WO-CA000915.
 PR 01-OCT-1998; 98US-00165264.

XX PA (BIOI-) BIO-ID DIAGNOSTIC INC.

XX DR PI; Vinayagamoorthy T;

XX PI; WPI; 2000-303800/26.

PT Nucleic acid analysis methods for simultaneously analyzing multiple nucleic acid regions for diagnosis and differentiation of pathological organisms comprises sequencing the nucleic acids in the reaction mixture.

PS Example 2; Page 23; 36pp; English.

This sequence represents a primer used in the analysis of a fragment of the bovine viral diarrhea virus (BVD) genome. The primer is used to illustrate the nucleic acid analysis methods of the invention. The methods are used for sequencing a nucleic acid in a mixture comprising two nucleic acid target sequences. The methods are used for simultaneously analysing multiple nucleic acid regions in a single reaction. This can allow the reliable diagnosis and differentiation of pathological organisms. The methods can be adapted to use a series of primers with additional sequences which allows the size of the amplified region to be increased. The technique is especially useful when the usual sequence of the region to be detected is known and the assay is being carried out to confirm its presence e.g. to rule out a falsely positive amplification reaction or to distinguish subsets of an organism of interest or allelic forms of a gene associated with a disease or predisposition to a disease. (Updated on 15-SEP-2003 to standardise OS

SQ sequence 320 BP; 4 A; 6 C; 308 G; 2 T; 0 U; 0 Other;

Query Match 18.6%; Score 97.4; DB 3; Length 320;
 Best Local Similarity 57.3%; Pred. No. 2.7e-09; Gaps 0;
 Matches 176; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

Qy 65 TCCACCGCCGACCATGAGTTCACCCAGTACGCCAACCTCTGAGGCC 124

RESULT 9
 AAZ77506
 ID AAZ77506 standard; cDNA; 2188 BP.

XX AC AAZ77506;
 XX PN DT 10-APR-2000 (first entry);
 XX DE Human ovarian tumor cDNA library derived EST fragment 57.
 XX KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;
 XX PR 09-APR-1998; 98DE-01017557.
 XX PA (META-) METAGEN GBS GENOMFORSCHUNG MBH.
 XX PI; Rosenthal A, Specht T, Hinermann B, Schmitt A, Pilarsky C, Dahl E;
 XX DR; WPI; 1999-551920/51.
 XX PS P-PSDB; AAY6638.

XX PT New nucleic acid sequences expressed in ovarian, and some other, cancer tissues, and derived polypeptides, for treatment of ovarian cancer and identification of therapeutic agents.

XX PS Claim 3; Page 184-185; 310pp; German.

XX CC This invention describes novel nucleic acid (cDNA) sequences (A) which have anticancer activity and are highly expressed in ovarian tumor tissue (and some also in testis and breast cancer tissue). The products of the invention can be used for gene therapy. (A) are used (i) for recombinant expression of polypeptides (B) and (ii) to isolate complete genes. (B) are used (i) to identify agents suitable for treatment of ovarian cancer; (ii) directly for treating this form of cancer (including expression from gene therapy vectors) and (iii) for generation of specific antibodies.

CC (A) are identified by assembling ESTs (expressed sequence tags) from a particular tissue type before comparison of expression patterns. This allows a significantly longer fragment of the gene to be revealed, so should reduce the number of failures associated with the fact that ESTs from different libraries may represent different parts of the same unknown gene, distorting the estimated frequency of occurrence in a

Sequence 2188 BP; 590 A; 567 C; 434 G; 597 T; 0 U; 0 Other;
 CC particular tissue. AAZ7450-Z7752 represent the human ovarian tumor cDNA library derived EST fragments described in the method of the invention and encode the protein fragments represented in AAY7605-Y76638

SQ

Query Match	18 6%: Score 97.2; DB 2; Length 2188;
Best Local Similarity	57 1%; Pred. No. 3e-09; Mismatches 177; Conservative 177; Matches 177; Conservatve 0; Mismatches 173; Indels 0; Gaps 0
Matches	177
Conservative	0
Mismatches	173
Indels	0
Gaps	0

QY

```

  66 CCACCGCCGCGGACCACTAGTTCACCCGTTACGCCACCTCTGACAGCCCCC 1
  5 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 6
  126 TTCTTCCCTTCCTAAAGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1
  65 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 1
  186 TCCCTTACCTGGCCATCTGGCTTGAGCTCCCTCTGCTACCCCTTTCATTC 2
  QY
  125 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 1
  246 CACCCCTTCCTCCACGCGGACCTAGCTCTCTCTCTCTCTCTCTCTCTCTCT 3
  QY
  185 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 2
  305 CTCTACGGAAGCGGTCTCCAACTGAACTGAACTGAACTGAACTGAACTGAA 3
  245 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 3
  QY
  366 CGCGCAGCA 375
  Db
  305 CCCCTCCCCA 314

```

RESULT 10

ID	AAA38184/C
ID	AAA38184 standard; DNA; 318 BP.
XX	
AC	
XX	
DT	15-SEP-2003 (revised)
DT	01-SEP-2000 (first entry)
DE	Primer used in the analysis of a BVDV genome fragment.
XX	
KW	Primer; bovine viral diarrhea virus; BVDV; nucleic acid analysis; diagnosis; pathological organism; detect; ss.
XX	
OS	Pestivirus type 1.
XX	
FN	WO2000020628-A1.
XX	
PD	13-APR-2000.
XX	
PF	01-OCT-1999; 99W0-CA000915.
XX	
PR	01-OCT-1998; 98US-00165264.
XX	
PA	(BIOI-) BIO-ID DIAGNOSTIC INC.
XX	
PI	Vinayagamoorthy T;
XX	
DR	WPI; 2000-303800/26.
XX	

PT Nucleic acid analysis methods for simultaneously analyzing multiple nucleic acid regions for diagnosis and differentiation of pathological organisms comprises sequencing the nucleic acids in the reaction mixture

PS Example 2; Page 23; 36pp; English.

CC This sequence represents a primer used in the analysis of a fragment of the bovine viral diarrhea virus (BVDV) genome. The primer is used to

CC

Search completed: April 14, 2004, 16:15:52
Total time: 438 seconds

New isolated PAL-18 polypeptide, useful for diagnosing, characterizing, and treating disease and in determining disease susceptibility.

Claim 1; Page 47-48, 150pp; English.

The present invention relates to human PAL-18 polypeptides and polynucleotides. The PAL-18 gene is found on chromosome 19q1. The sequences can be used to diagnose, monitor and treat cancers, particularly breast, colon and prostate cancers. The present sequence is a PAL-18 polynucleotide shown in the invention.

Sequence 1064 BP; 93 A; 601 C; 51 G; 194 T; 0 U; 125 Other;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

On nucleic - nucleic search, using sw model

Run on: April 14, 2004, 14:27:16 ; Search time 97 Seconds
(without alignments)
2992.158 Million cell updates/sec

Title: US-09-669-817A-4
Perfect score: 523

Sequence: 1 atctgatggatcgatcacca.....gaccgcgaatggaa 523

Scoring table: IDENTITY NUC
Gapov 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Maximum DB seq length: 0

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/B.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/C.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/backfile1.seq: *
6: /cgn2_6/ptodata/2/ina/backfile2.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Length DB ID Description

Result No.	Score	Query Length	DB	ID	Description
1	109.2	20.9	7218	1	US-09-232-463-14
C	2	98	18.7	320	US-09-165-264-14
C	3	97.4	18.6	320	US-09-165-264-13
C	4	97	18.5	319	US-09-165-264-8
C	5	95.8	18.3	320	US-09-165-264-7
C	6	95.4	18.2	318	US-09-165-264-12
C	7	95.2	18.2	320	US-09-165-264-11
C	8	81.2	15.5	320	US-09-249-55A-4
C	9	81.1	15.5	1926	US-09-249-55A-4
C	10	78	14.9	5326	US-08-136-2
C	11	78	14.9	53577	US-08-158-136-1
C	12	76.6	14.6	4466	US-09-410-551B-20
C	13	76.6	14.6	4478	US-09-410-551B-16
C	14	76.6	14.6	4547	US-09-410-551B-22
C	15	76.6	14.6	4571	US-09-410-551B-18
C	16	76.6	14.6	77536	US-09-410-551B-1
C	17	76.4	14.6	12001	US-08-458-168A-11
C	18	74.4	14.2	12331	US-09-128-155-16
C	19	73.8	14.1	1926	US-09-249-55A-2
C	20	73.8	14.1	1926	US-09-410-399-3
C	21	73.8	14.1	2580	US-09-410-399-2
C	22	73.8	14.1	5452	US-09-359-081-2
C	23	73.8	14.1	8705	US-09-130-14-1
C	24	73.8	14.1	9600	US-09-910-647-1
C	25	73.8	14.1	9600	US-09-620-945-1
C	26	73.8	14.1	10596	US-07-884-811-15

ALIGNMENTS

C	28	73.8	14.1	10596	1	US-07-885-971-15	Sequence 15, Appl
C	29	73.8	14.1	10596	1	US-08-087-782A-15	Sequence 15, Appl
C	30	73.8	14.1	10596	1	US-08-194-088B-15	Sequence 15, Appl
C	31	73.8	14.1	10596	2	PCT-US91-0468-15	Sequence 15, Appl
C	32	73.8	14.1	10596	5	US-09-224-562A-48	Sequence 48, Appl
C	33	73.8	14.1	10600	4	US-09-410-551B-1	Sequence 1, Appl
C	34	73.6	14.1	7736	4	US-09-410-551B-1	Patent No. 519616
C	35	72.6	13.9	4897	6	519616-7	Sequence 209, App
C	36	71.4	13.7	51259	3	US-08-731-891-209	Sequence 8385, AP
C	37	71.4	13.7	51259	4	US-09-618-165-209	Sequence 8423, AP
C	38	65.4	12.5	925	3	US-08-858-003-1	Sequence 1, Appl
C	39	65.4	12.5	925	3	US-09-078-166-1	Sequence 1, Appl
C	40	65.4	12.5	925	3	US-08-937-467-1	Sequence 1, Appl
C	41	64.4	12.3	438	4	US-09-232-91A-8422	Sequence 8422, AP
C	42	64.4	12.3	471	4	US-09-232-991A-8385	Sequence 8385, AP
C	43	64.4	12.3	1751	1	US-09-252-991A-8423	Sequence 1, Appl
C	44	63.8	12.2	31571	1	US-08-323-433B-1	Sequence 2, Appl
C	45	63.2	12.1	4403765	3	US-09-103-840A-2	Sequence 2, Appl

RESULT 1	US-08-232-463-14	Sequence 14 Application US/08232463	Patent No. 5670367
		GENERAL INFORMATION:	
		APPLICANT: DORNER, F.	
		APPLICANT: SCHEPFLINGER, F.	
		APPLICANT: FALKNER, F. G.	
		TITLE OF INVENTION: RECOMBINANT FOWLPX VIRUS	
		NUMBER OF SEQUENCES: 52	
		CORRESPONDENCE ADDRESS:	
		ADDRESSEE: Foley & Lardner	
		STREET: 1800 Diagonal Road, Suite 500	
		CITY: Alexandria	
		STATE: VA	
		COUNTRY: USA	
		ZIP: 22313-0299	
		COMPUTER READABLE FORM:	
		COMPUTER: IBM PC compatible	
		OPERATING SYSTEM: PC-DOS/MS-DOS	
		SOFTWARE: Patentin Release #1.0, Version #1.25	
		CURRENT APPLICATION DATA:	
		APPLICATION NUMBER: US/08/232,463	
		FILING DATE:	
		CLASSIFICATION: 435	
		PRIORITY APPLICATION DATA:	
		APPLICATION NUMBER: US/07/935,313	
		FILING DATE:	
		APPLICATION NUMBER: EP 91 114 300.6	
		FILING DATE: 26-AUG-1991	
		ATTORNEY/AGENT INFORMATION:	
		NAME: BENT, Stephen A.	
		REGISTRATION NUMBER: 29,768	
		REFERENCE/DOCKET NUMBER: 30-72/114 IMMU	
		TELECOMMUNICATION INFORMATION:	
		TELEPHONE: (703) 836-9300	
		TELEFAX: (703) 683-4109	
		TELEX: 839149	
		INFORMATION FOR SEQ ID NO: 14:	
		SEQUENCE CHARACTERISTICS:	
		LENGTH: 7218 base pairs	
		REFERENCE/DOCKET NUMBER: 30-72/114 IMMU	
		TYPE: nucleic acid	
		STRANDEDNESS: single	
		TOPOLOGY: linear	
		IMMEDIATE SOURCE: CLONE: PTZP-F1s	
		US 08-232-463-14	
Query Match:	20.9%	Score 109.2;	DB 1; Length 7218;

; FEATURE: OTHER INFORMATION: Description of Artificial Sequence:Primer sequence

; US-09-165-264-8

Query Match 18.5%; Score 97; DB 3; Length 319;
Best Local Similarity 56.9%; Pred. No. 1.8e-11; Matches 178; Conservative 0; Mismatches 135; Indels 0; gaps 0;

Qy 58 AGAGACTCCACCGCCGGGGCACCATGAGTCTCACACGGTTAGCCAACTCTCG 117
Db 313 Aaggctgcacaccccccaccccccaccccccaccccccaccccccacccccc 254

Qy 118 ACCACCCCTCTCCCTTCACCCGGCAGCTCTCTCCACCCCTCTCCAC 177
Db 253 ccc 194

Qy 178 CCGCGTTCTCACTGGCCATTCTGGCTAGACTCCCTTGCTGACCTT 237
Db 193 ccc 134

Qy 238 TCCATCTCACCCCTCTCCACGGACACTASCTCCCTCTGACCCCTCTCC 297
Db 133 ccc 74

Qy 298 TCCACACCTCACGCCGGTCTCCAACTCGAGCTGGACCCACACCCAC 357
Db 73 cc 14

Qy 358 ACCCAGGGAC 370
Db 13 cccccccccccc 1

RESULT 5 US-09-165-264-7/c

; Sequence 7, Application US/09165264
; Patent No. 619510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamoorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/165,264
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence

US-09-165-264-12

Query Match 18.2%; Score 95.4; DB 3; Length 318;
Best Local Similarity 57.0%; Pred. No. 3.8e-11; Matches 174; Conservative 0; Mismatches 131; Indels 0; gaps 0;

Qy 64 CTCAACGGCGGGGACGACCTAGAGTCTCACACCTTAACCCACACTCTGGACAC 123
Db 306 CACCAACCC 247

Qy 124 CCTCTCTCCCTCTCCACCCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 183
Db 246 CCTCC 187

Qy 184 TTCTCTCACCTCCCTCTGGCTCTGAATCTCTCTCTGTCACCTTTCATC 243
Db 186 CCACCC 127

Qy 244 TCCACCCCTCTCCACGGACACTAGTCACCTCTCACACCCCTCTCTCCACA 303
Db 126 CC 67

Qy 304 CCTCACCGACGGAGCTGGCTCTCTGGCTAGCTGGCTACCCCTGGACACCCA 363
Db 66 CCC 7

Qy 364 Gccca 368
Db 6 ccccc 2

Query Match 18.3%; Score 95.8; DB 3; Length 320;
Best Local Similarity 57.5%; Pred. No. 3.2e-11; Matches 172; Conservative 0; Mismatches 127; Indels 0; gaps 0;

Qy 54 CTCCACCCGGGGCACCATGAGTCTCACACGGTTACCCACCTCTCGACGCC 123
Db 302 CTCC 243

Qy 124 CCTCTCTCACCTCACCGCCGCTCTCTCTGGACCTCTCTCCACCCGGCG 183
Db 242 CCC 183

Qy 184 TTTCTCCACCTCCATTTGGCCCTGACTCCCTCTGGCTGACCCCTTCCATC 243
Db 182 CCC 123

Qy 244 TCCACCCCTCTCACCGACGGACACTAGTCACCTCTCACACCCCTCTCCACA 303
Db 122 CCC 63

RESULT 6 US-09-165-264-11/c

; Sequence 11, Application US/09165264
; Patent No. 619510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamoorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence

US-09-165-264-12

Query Match 18.2%; Score 95.4; DB 3; Length 318;
Best Local Similarity 57.0%; Pred. No. 3.8e-11; Matches 174; Conservative 0; Mismatches 131; Indels 0; gaps 0;

Qy 64 CTCAACGGCGGGGACGACCTAGAGTCTCACACCTTAACCCACACTCTGGACAC 123
Db 306 CACCAACCC 247

Qy 124 CCTCTCTCCCTCTCCACCCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 183
Db 246 CCTCC 187

Qy 184 TTCTCTCACCTCCCTCTGGCTCTGAATCTCTCTCTGTCACCTTTCATC 243
Db 186 CCACCC 127

Qy 244 TCCACCCCTCTCCACGGACACTAGTCACCTCTCACACCCCTCTCTCCACA 303
Db 126 CC 67

Qy 304 CCTCACCGACGGAGCTGGCTCTCTGGCTAGCTGGCTACCCCTGGACACCCA 363
Db 66 CCC 7

Qy 364 Gccca 368
Db 6 ccccc 2

RESULT 7 US-09-165-264-11/c

; Sequence 11, Application US/09165264
; Patent No. 619510
; GENERAL INFORMATION:
; APPLICANT: Vinayagamoorthy, Thuraiayah
; TITLE OF INVENTION: Multi-Loci Genomic Analysis
; FILE REFERENCE: 44747
; CURRENT APPLICATION NUMBER: US/09/165,264
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Primer sequence


```

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 4466
; OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
; FEATURE: OTHER INFORMATION: PKS synthase fragment
; ORGANISM: Artificial Sequence
; LOCATION: (9)..(4454)
; US-09-410-551B-20

Query Match 14.6%; Score 76 6; DB 4; Length 4466;
Best Local Similarity 53.4%; Pred. No. 3.2e-07;
Matches 183; Conservative 0; Mismatches 159; Indels 1; Gaps 1;

Qy 32 CACTCCACCTCAGGGACTGAGAGTCAACGCCGACCATGAGTC 91
Db 3741 CACCCCAACCCCTCCCTGGCGCACATGCCACCTGACCC 3800
Qy 92 TCACCCACGTACGCCAACCTCTGGACACCCCTTCACCCGGTCTC 151
Db 3801 -CTAACCCACACCTCACACCCACCTCACCCCTCACACCCAC 3859
Qy 152 CTCCCTCTACCCCTCTCTCCACCCGGTTCCACATGCCATTTCGCT 211
Db 3860 CACCAACACCCCTCAACCCGAACACCCATCATCATCACGGGACTCCGGACCT 3919
Qy 212 CGACTCCCTTCGCTGCAGACCTTTCATCCACCCCTCTCCACCCGGACCC 271
Db 3920 CGCCGGCATCTCGGCCAACCTGAAACCCACACCTACTCTCCGGACCC 3979
Qy 272 TAGCTCCCTCTCACCCCTCTGGACACCCCTCACCGACGGCTCAACTGA 331
Db 3980 ACCCCGAGACCCACCCGGACACTCTGAGGGAGACCCACACT 4039
Qy 4040 CGCCACCACTCACCCACATCCCCAACCCCTCACCGCATC 4094

RESULT 13
US-09-410-551B-16
Sequence 16, Application US/09410551B
Patent No. 6503737
GENERAL INFORMATION:
APPLICANT: KOSAN BIOSCIENCES, Inc.
APPLICANT: REEVES, CHRISTOPHER
APPLICANT: CHU, DANIEL
APPLICANT: KHOSLA, CHAITAN
APPLICANT: SANTI, DANIEL
APPLICANT: WU, KAI
TITLE OF INVENTION: POLYPEPTIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
FILE REFERENCE: 30062-20026.00
CURRENT FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: US/09/410,551B
PRIOR FILING DATE: 1999-06-17
PRIOR APPLICATION NUMBER: US 60/123,810
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: US 60/102,748
PRIOR FILING DATE: 1998-10-02
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 16
LENGTH: 4478
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
OTHER INFORMATION: PKS synthase fragment
OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
OTHER INFORMATION: PKS synthase fragment
OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
OTHER INFORMATION: PKS synthase fragment
; US-09-410-551B-22

Query Match 14.6%; Score 76 6; DB 4; Length 4547;
Best Local Similarity 53.4%; Pred. No. 3.2e-07;
Matches 183; Conservative 0; Mismatches 159; Indels 1; Gaps 1;

Qy 32 CACTCCACCTCAGGGACTGAGAGTCAACGCCGACCATGAGTC 91
Db 3753 CACCCCAACCCCTCCCTGGACACCTGACCC 3812
Qy 92 TCACCCACGTACGCCAACCTCTGGACACCCCTTCACCCGGTCTC 151
Db 3813 -CTAACCCACACCTCACACCCACCTCACCCCTCACACCCAC 3871
Qy 152 CGACTCCCTTCGCTGCAGACCTTTCATCCACCCCTCTCCACCCGGACCC 271
Db 3932 CGCCGGCATCTCGCCGACCTGAAACCCACCTACTCTCCGGACCC 3991
Qy 272 TAGCTCCCTCTCACCCCTCTGGACACCCCTCACCGACGGCTCAACTGA 331
Db 3992 ACCCCGAGACCCACCCGGACCCACCTCTGGAGCTGGAGTCGGACCCACACT 4051
Qy 332 GCTCCCTCTGGCCACGGCCACCCGGCCACCCGGCCACAGC 374
Db 4040 CGCCACCACTCACCCACATCCCCAACCCCTCACGGCATC 4094

RESULT 14
US-09-410-551B-22
Sequence 22, Application US/09410551B
Patent No. 6503737
GENERAL INFORMATION:
APPLICANT: KOSAN BIOSCIENCES, Inc.
APPLICANT: REEVES, CHRISTOPHER
APPLICANT: CHU, DANIEL
APPLICANT: KHOSLA, CHAITAN
APPLICANT: SANTI, DANIEL
APPLICANT: WU, KAI
TITLE OF INVENTION: POLYPEPTIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
FILE REFERENCE: 30062-20026.00
CURRENT FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: US/09/410,551B
PRIOR FILING DATE: 1999-06-17
PRIOR APPLICATION NUMBER: US 60/123,810
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: US 60/102,748
PRIOR FILING DATE: 1998-10-02
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 22
LENGTH: 4547
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
OTHER INFORMATION: PKS synthase fragment
FEATURE: OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
OTHER INFORMATION: PKS synthase fragment
FEATURE: OTHER INFORMATION: Description of Artificial Sequence: DNA encoding synthetic
OTHER INFORMATION: PKS synthase fragment
; US-09-410-551B-22

```


OM nucleic - nucleic search, using sw model
Copyright (c) 1993 - 2004 Compugen Ltd.

```

17: /egn2_6/ptodata/2/pubnra/USO_NRM_PUB.seq:*
18: /egn2_6/ptodata/2/pubnra/USO6_PUBOMB.seq:*

```

```

C 16 91.8 3133 14 US-10-017-161-1483 Sequence 1483, AP
C 17 91.8 3133 15 US-10-017-161-1483 Sequence 1483, AP
C 17.5 91.6 3133 12 US-10-024-591-104895 Sequence 104895,
C 18 90.2 956 12 US-10-027-632-31508 Sequence 31508, A
C 19 90.2 956 12 US-10-027-632-31508 Sequence 31508, A
C 20 90.2 17.2 956 12 US-10-027-632-31508 Sequence 31508, A
C 21 89.4 17.1 1117 14 US-10-017-161-1403 Sequence 1403, AP
C 22 89.4 17.1 1117 15 US-10-222-798-1141 Sequence 1141, AP
C 23 88.8 17.0 485 12 US-10-424-599-65670 Sequence 65670, A
C 24 87.8 16.8 931 14 US-10-198-846-8585 Sequence 8585, AP
C 25 86.2 16.5 492 12 US-10-424-599-59619 Sequence 59619, A
C 26 84.2 16.1 636 12 US-10-424-599-59619 Sequence 59619, A
C 27 84.2 16.1 1111 12 US-10-425-114-32750 Sequence 32750, A
C 28 83.8 16.0 1327 15 US-10-398-221-1265 Sequence 1265, AP
C 29 83.4 15.9 805 12 US-10-424-599-61903 Sequence 61903, A
C 30 83.4 15.9 809 12 US-10-424-599-26988 Sequence 26988, A
C 31 83 15.9 1009 12 US-10-424-599-1962 Sequence 1962, AP
C 32 83 15.9 1087 12 US-10-424-599-1962 Sequence 1962, AP
C 33 82.4 15.8 791 12 US-10-424-599-64474 Sequence 8404, A
C 34 82.4 15.7 883 12 US-10-027-632-4358 Sequence 4358, AP
C 35 82.2 15.7 883 15 US-10-027-632-4358 Sequence 4358, AP
C 36 81.6 15.6 1788 12 US-10-027-632-97688 Sequence 97688, A
C 37 81.6 15.6 1788 12 US-10-027-632-100372 Sequence 100372, A
C 38 81.6 15.6 1788 15 US-10-027-632-97688 Sequence 97688, A
C 39 81.6 15.6 1788 15 US-10-027-632-100372 Sequence 100372, A
C 40 81.4 15.6 824 12 US-10-424-599-88450 Sequence 88450, A
C 41 81.2 15.5 863 12 US-10-027-632-31636 Sequence 31636, A
C 42 81.2 15.5 863 15 US-10-027-632-31636 Sequence 31636, A
C 43 81.2 15.5 922 14 US-017-161-1435 Sequence 1435, AP
C 44 81.2 15.5 922 15 US-10-222-798-1161 Sequence 1151, AP
C 45 81.2 15.5 922 15 US-10-424-599-139811 Sequence 139811, AP

ALIGNMENTS

RESULT 1
US-10-425-114-19216
; Sequence 19216, Application US/10025114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21 (53313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 19216
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE: OTHER INFORMATION: Clone ID: LIB3137-017-DB_FLI
US-10-425-114-19216
Query Match 24.9%; Score 130.4; DB 12; Length 1454;
Best Local Similarity 59.3%; Pred. No. 8-1e-23; Matches 305; Conservative 0; Mismatches 186; Indels 23; Gaps 4;
QY 13 GATCACCAGTCCTCACATCCACGCCGGAGCTGAGAGCTCACCAC 72
Db 46 GATGCCATTCACCGCTCCGACTCCACCTCTGAGACCCCTCTTCC 72
QY 73 CGCGGCCACGAGTAGCCACCTCTGAGACCCCTCTTCC 1105
Db 106 CCTTCCTTACCCGTGCTGCGCTGCGCTACGCGCT-CCTGGCCACAC 163
Ov 133 CCTTCCTTACCCGTGCTGCGCTGCGCTACGCGCT-CCTGGCCACAC 192

```


Sequence 47, Application US/10032393
 Publication No. US20030027286A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Wall, Daniel
 APPLICANT: Gross, Molly
 TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
 FILE REFERENCE: ELITRA_010A
 CURRENT APPLICATION NUMBER: US/10/032,393
 CURRENT FILING DATE: 2001-12-21
 PRIOR APPLICATION NUMBER: 60/259,434
 PRIOR FILING DATE: 2000-12-27
 PRIOR APPLICATION NUMBER: 09/948,993
 PRIOR FILING DATE: 2001-09-06
 NUMBER OF SEQ ID NOS: 68
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 47
 LENGTH: 12733
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Vector PEPEF1
 US-10-032-393-47

Query Match 19.5%; Score 102; DB 14; Length 12733;
 Best Local Similarity 54.9%; Pred. No. 6.3e-16;
 Matches 201; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
 QY 6 ATCAGGAGATGCCAACGTCCTCACACTCCACATGGCAAGGAACTGAGGAGAGCT 65
 Db 5357 ATGGGTGACCC 5298
 QY 66 CCACCGCCGCCGCCACATGAGTCTCACACCGGTACGCCACCTCTGAGGACCC 125
 Db 5297 CCC 5338
 QY 126 TTCTTCTCTTACACCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 185
 Db 5237 CCCCCCCCCAC 5178
 QY 186 TCT 245
 Db 5177 CCCCCCCCCAC 5118
 QY 246 CACCCCTCTCCAGCGGACACTAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 305
 Db 5123 CCCCCCCCCAC 5064
 QY 246 CACCCCTCTCCAGCGGACACTAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 305
 Db 5063 CCCCCCCCCAC 5004
 QY 366 CCCACC 371
 Db 5003 CCCCCC 4998

RESULT 5
 US-10-032-393-8/C
 Sequence 8, Application US/10032393
 Publication No. US20030027286A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Wall, Daniel
 APPLICANT: Gross, Molly
 TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
 FILE REFERENCE: ELITRA_010A
 CURRENT APPLICATION NUMBER: US/10/032,393
 CURRENT FILING DATE: 2001-12-21
 PRIOR APPLICATION NUMBER: 60/259,434
 PRIOR FILING DATE: 2000-12-27

PRIOR APPLICATION NUMBER: 09/948,993
 PRIOR FILING DATE: 2001-09-06
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 NUMBER OF SEQ ID NOS: 68
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 8
 LENGTH: 12739
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Vector PEPEF1
 US-10-032-393-8

Query Match 19.5%; Score 102; DB 14; Length 12739;
 Best Local Similarity 54.9%; Pred. No. 6.3e-16;
 Matches 201; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
 QY 6 ATCAGCAGATGCCAACGTCCTCACACTCCACATGGCAAGGAACTGAGGAGAGCT 65
 Db 5363 ATCGTCGACCC 5304
 QY 66 CCACCGCCGCCGCCACATGAGTCTCACACCGGTACGCCACCTCTGAGGACCC 125
 Db 5303 CCCCCCCCCAC 5244
 QY 126 TTCTTCTCTTACACCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 185
 Db 5243 CCCCCCCCCAC 5184
 QY 186 TCT 245
 Db 513 CCCCCCCCCAC 5124
 QY 246 CACCCCTCTCCAGCGGACACTAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 305
 Db 5123 CCCCCCCCCAC 5064
 QY 366 CCCACC 371
 Db 5003 CCCCCC 4998

RESULT 6
 US-10-017-161-1857/C
 Sequence 1857, Application US/10017161
 Publication No. US20030143668A1
 GENERAL INFORMATION:
 APPLICANT: SUMI, MAKIKO
 APPLICANT: ASAI, KIYOSHI
 APPLICANT: AKIYAMA, HIROYUKI
 APPLICANT: ABURATANI, YUJAKA
 TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
 FILE REFERENCE: 08A335/0152
 CURRENT APPLICATION NUMBER: US/10/017,161
 CURRENT FILING DATE: 2002-12-18
 PRIOR APPLICATION NUMBER: JP 2001/246789
 PRIOR FILING DATE: 2001-06-18
 NUMBER OF SEQ ID NOS: 2430
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1857
 LENGTH: 3163
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: source
 LOCATION: (1)..(3163)
 FEATURE:
 NAME/KEY: CDS

LOCATION: (201)..(2963)
FEATURE:
NAME/KEY: modified_base
LOCATION: (658)..(669)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (671)..(677)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (704)..(707)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (708)..(736)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (739)..(743)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (745)..(747)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (750)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (754)..(756)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (755)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (756)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (761)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (818)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (892)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (920)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (897)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (941)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (955)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1082)..(1084)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1105)..(1107)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1130)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1147)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1162)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1202)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1233)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1247)..(1248)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1485)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1505)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1601)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified_base
LOCATION: (1635)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:

NAME/KEY: modified_base

LOCATION: (1637)

OTHER INFORMATION: a, t, c, g, unknown or other

FEATURE: modified_base

NAME/KEY: modified_base

LOCATION: (1654)

OTHER INFORMATION: a, t, c, g, unknown or other

FEATURE: modified_base

NAME/KEY: modified_base

LOCATION: (1674)

OTHER INFORMATION: a, t, c, g, unknown or other

FEATURE: modified_base

NAME/KEY: modified_base

LOCATION: (1695)

OTHER INFORMATION: a, t, c, g, unknown or other

FEATURE: modified_base

NAME/KEY: modified_base

LOCATION: (1866)..(1965)

OTHER INFORMATION: a, t, c, g, unknown or other

FEATURE: modified_base

NAME/KEY: modified_base

LOCATION: (1852)

OTHER INFORMATION: a, t, c, g, unknown or other

FEATURE: modified_base

NAME/KEY: modified_base

LOCATION: (1968)..(1980)

OTHER INFORMATION: a, t, c, g, unknown or other

FEATURE: modified_base

NAME/KEY: modified_base

LOCATION: (1982)..(2009)

Query Match 19.4%; Score 101.4; DB 14; Length 3163;
 Best Local Similarity 51.1%; Pred. No. 9.7e-16; Pairs 0; Mismatches 198; Indels 0; Gaps 0;
 Matches 207; Conservative 0; Other Information: a, t, c, g, unknown or other
 QY 6 ATCAGCGATCCCAAGCTCTTCACTCCACTCCAGGGACTGAGGAGAGT 65
 Db 1270 AACACCCCNCCGAACCCCNNCCCCCCCCCNCGGCCCGGAGCCCCCCC 1211
 QY 66 CACCGGAGGGGACCATGAGTCACCCGGTACGCCACCTCTGAGCACCCCC 125
 Db 1210 CCCCCCNCACCC 1151
 QY 126 TGTTCCTTCCACCGGGCTCTCTCTCTGCCCCCTCTGAGCACCCCCGGTT 185
 Db 1150 CCNCC 1091
 QY 186 TCTCCACTGCCATCTCCCTGACTGACCCCCCTGCTGCTGACCCCTTCATCTC 245
 Db 1090 CCCCCNNCC 1031
 QY 246 CACCCCTCTCTCCACGCCAACTAGCTCTCTCTGACCCCTCCCTCACACC 305
 Db 1030 CCC 971
 QY 305 CTGACCGACGGCTCTCAACTGACTGCCCTGACCCGACCCAGG 365
 Db 970 CCC 911
 QY 366 CTCACCGAGGAAGTCACCTATGACCCGACCCGGCG 410
 Db 910 CCC 866

RESULT 7 US-10-292-798-1513/c
 Sequence 1513, Application US/10292798
 Publication No. US2003023583A1
 GENERAL INFORMATION:
 APPLICANT: SWWA, MAKIKO
 APPLICANT: ASAI, KIYOKI
 APPLICANT: AKIYAMA, YUTAKA
 APPLICANT: ABURATANI, HIROVIKI

TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
 FILE REFERENCE: 084335/166
 CURRENT APPLICATION NUMBER: US/10/292-798
 CURRENT FILING DATE: 2002-11-13
 PRIOR APPLICATION NUMBER: 10/017,161
 PRIOR FILING DATE: 2001-12-18
 PRIOR APPLICATION NUMBER: JP 2001-246789
 NUMBER OF SEQ ID NOS: 2070
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 1513
 LENGTH: 3163
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE: source
 FEATURE: (1)..(3163)
 LOCATION: (201)..(2963)
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (556)..(655)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (658)..(669)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (671)..(697)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (709)..(736)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (701)..(707)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (739)..(743)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (745)..(747)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (750)..(750)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (754)..(756)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (761)..(761)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (763)..(763)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE: modified_base
 NAME/KEY: modified_base
 LOCATION: (818)..(818)
 OTHER INFORMATION: a, t, c, g, unknown or other

RESULT 8

US-10-017-161-1481/c
 Sequence 1481, Application US/10017161
 Publication No. US/0030143668A1

GENERAL INFORMATION:
 APPLICANT: SWIA, MAKIKO
 APPLICANT: ASAI, KIYOSHI
 APPLICANT: ARIYAMA, YUTAKA
 APPLICANT: ABURATANI, HIROKU
 TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
 FILE REFERENCE: 084335/0152
 CURRENT APPLICATION NUMBER: US/0/017,161
 CURRENT FILING DATE: 2002-12-18
 PRIOR APPLICATION NUMBER: JP 2001/246789
 PRIOR FILING DATE: 2001-06-18
 NUMBER OF SEQ ID NOS: 2430
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 1481
 LENGTH: 5452
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 LOCATION: SOURCE
 FEATURE:
 LOCATION: (1) .. (5452)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (201) .. (5252)
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (11)..(140)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (145)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (147)..(148)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (149)..(150)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (151)..(152)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (153)..(154)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (155)..(156)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (157)..(158)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (159)..(160)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (161)..(162)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (163)..(164)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (165)..(166)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (167)..(168)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (169)..(170)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (171)..(172)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (173)..(174)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (175)..(176)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (177)..(178)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (179)..(180)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (181)..(182)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (183)..(184)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (185)..(186)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (187)..(188)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (189)..(190)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (191)..(192)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (193)..(194)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (195)..(196)

LOCATION: (162)..(162)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (165)..(165)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (168)..(168)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (189)..(189)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (191)..(191)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (199)..(199)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (196)..(196)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (214)..(214)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (221)..(221)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (218)..(218)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (225)..(225)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (232)..(232)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (229)..(229)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (246)..(246)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (256)..(256)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (248)..(248)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (261)..(261)

OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (271)..(273)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (283)..(285)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (276)..(277)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (296)..(296)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (302)..(302)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (298)..(300)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (307)..(307)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (314)..(315)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (339)..(339)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (343)..(343)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (345)..(348)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (351)..(352)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (357)..(357)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (359)..(359)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (362)..(362)
 OTHER INFORMATION: a, t, c, g, unknown or other
 FEATURE:
 NAME/KEY: modified base
 LOCATION: (365)..(365)
 OTHER INFORMATION: a, t, c, g, unknown or other

US-10-424-599-54839
Sequence 54839, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovacic David R
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21-(53223)B
CURRENT APPLICATION NUMBER: US/10-424, 599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 283684

RESULT 12
US-10-424-599-2320
; Sequence 2320, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yinhua
; APPLICANT: Cao Yongwei
; LENGTH: 778
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(778)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20531C.1
US-10-424-599-5839

TITLE OF INVENTION: SoY Nucleic Acid Molecules and Oligo Molecules Associated with Plants and Uses Thereof for Plant Improvement.
FILE REFERENCE: 38-21(5323)B
CURRENT APPLICATION NUMBER: US/10/424,599
CIPRIMENT FILING DATE: 2013-04-29

RESULT 14
 US-09-804-682-33
 ; Sequence 33, Application US/09804682
 ; Patent No. US20020106765A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kinders, Robert
 ; APPLICANT: Corey, Michael J.
 ; TITLE OF INVENTION: PVAP-18 POLYPEPTIDE
 ; TITLE OF INVENTION: ENCODING THE SAME
 ; TITLE OF INVENTION: MODULATING THE
 FILE REFERENCE: 130001_406

SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 33
 LENGTH: 1065
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 14, 21, 27, 33, 36, 42, 72, 107,
 LOCATION: 227, 235, 240, 242, 243, 247,
 LOCATION: 26, 321, 322, 324, 330, 332,
 LOCATION: 390, 393, 399, 401, 407, 415,
 OTHER INFORMATION: n = A, T, C or G
 NAME/KEY: misc_feature
 LOCATION: 479, 494, 501, 508, 511, 513, 5
 LOCATION: 623, 624, 633, 653, 654, 656, 6
 LOCATION: 721, 722, 735, 743, 747, 748,
 LOCATION: 798, 802, 808, 810, 813, 814, 8
 OTHER INFORMATION: n = A, T, C or G
 NAME/KEY: misc_feature
 LOCATION: 842, 847, 888, 900, 909, 910, 9
 LOCATION: 973, 1002, 1005, 1028, 1029, 10
 OTHER INFORMATION: n = A, T, C or G
 US-09-804-682-33

1, 103, 207, 208, 221, 223,
248, 259, 263, 269, 273, 278,
335, 336, 340, 367, 371, 385,

Homo sapiens	604
misc feature	
55 - 167,	21, 24, 25, 33, 39,
235, 237,	238, 244, 245, 246,
321, 323,	330, 334, 340, 349,
397, 405,	432, 437, 454, 455,
FORMATION:	n = A, T, C or G
misc feature	
485 - 487,	488, 494, 496, 499,
583, 600,	611, 613, 623, 624,
694, 701,	713, 716, 720, 721,
744, 781,	782, 785, 789, 799,
FORMATION:	n = A, T, C or G
MISC feature	
858 - 878,	884, 886, 896, 897,
957, 961,	965, 981, 991, 993,
1043,	1047, 1049, 1051,
FORMATION:	1054, n = A, T, C or G

9, 214, 231, 232,
92, 297, 306, 319,
63, 372, 376, 378,
59, 468, 470

1, 48, 103, 207, 208, 221, 223,
 248, 259, 259, 263, 269, 273, 278,
 335, 335, 336, 340, 367, 371, 385,
 421, 447, 454, 471, 475
 425, 539, 540, 579, 595, 605,
 658, 664, 670, 679, 711, 713,
 754, 776, 778, 779, 780, 783,
 820, 822, 824, 825, 838
 916, 926, 927, 943, 948, 962,
 1034, 1057, 1065

; DB 9; Length 1065;
 3.5e-14; Indels 0; Gaps:
 198;

Homo sapiens	
misc. feature	
5, 6, 16, 21, 24, 25, 33, 39,	72, 110, 20,
235, 237, 238, 244, 245, 246,	256, 282, 2,
321, 323, 330, 334, 340, 349,	354, 355, 3,
397, 405, 432, 454, 454,	455, 457, 458,
FORMATION: n = A, T, C or G	4
misc. feature	
485, 487, 488, 494, 496, 499,	511, 524, 5
583, 600, 611, 613, 623,	634, 652, 654,
694, 701, 713, 716, 720,	721, 725, 731,
744, 781, 782, 785, 789,	799, 803, 821,
858, 878, 884, 886, 896,	897, 901, 917, 9
957, 961, 965, 981, 991,	993, 1001, 1002,
1043, 1047, 1049, 1051,	1054, 1056
FORMATION: n = A, T, C or G	-29
Similarity	17.6%;
9; Conservative	Score 92.2; DB 9;
	Pred. No. 1. 1.e-13; Mismatches 171;
CACTCCACTCAGGGAGCTGAGAGAGTCCACGGCC	
cccccccccncnccctcnccnctc	
TCACCCCGTAGGCCACCTCTCGAACCCCTCTCTCC	
cacccatctacccctca	
ctctctcgacccttc	
cctcccccnncccttc	
cgactcccccrttgatcgatcgatccatccacccctt	
NONCCCCCCCCTCC	
TAGCTCCTCTGACCCCTCTCTCACACCCCTCGAC	
TNCNCNCNCTCNCCNCCNCCNCCNCTTC	
GCTGCCTCGACGGCCACACCCAGGCCAC	
TCCCCCANNCCNTTCCCCCNCCNCCCCCCCC	
5 secs	dated: April 14, 2004, 14:27:09

```

LENGTH: 1064
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE: misc_feature
NAME/KEY: misc-feature
LOCATION: 5, 6-, 21, 24, 25, 33, 39, 72, 110, 209, 214, 233,
LOCATION: 235, 237, 238, 244, 245, 246, 256, 282, 292, 297, 311,
LOCATION: 321, 323, 330, 334, 340, 349, 354, 355, 363, 372, 381,
LOCATION: 397, 405, 432, 437, 454, 455, 457, 458, 459, 468,
OTHER INFORMATION: n = A, T, C or G
NAME/KEY: misc-feature
LOCATION: 485, 487, 488, 494, 496, 498, 511, 524, 527, 552,
LOCATION: 583, 600, 611, 613, 623, 624, 652, 654, 674, 691, 699,
LOCATION: 694, 701, 713, 716, 720, 721, 725, 731, 734, 735,
LOCATION: 744, 782, 785, 799, 803, 821, 823, 847, 851,
OTHER INFORMATION: n = A, T, C or G
NAME/KEY: misc-feature
LOCATION: 858, 878, 884, 886, 896, 897, 901, 917, 926, 932, 948,
LOCATION: 957, 961, 965, 981, 991, 993, 1001, 1002, 1005, 1011,
LOCATION: 1043, 1047, 1049, 1051, 1054, 1056
OTHER INFORMATION: n = A, T, C or G
US-09-804-682-29

Query Match                                17.6%; Score 92.2; Length 1064
Best Local Similarity 49.7%; pred. No. 1.9e-13; DB 9; Mismatches 171; Indels 8
Matches 169; Conservative 0; Mismatches 171; Indels 8
Qy      32 CACTCCACCTCAGGGGAGCTGAGAGAGCTTCAACCGCCGCCA
Db      472 CCCCTCCCCCNCCNCNCCNTNCNTCTCCCCCNCCCTCCCCC
Qy      92 TCACCCCGTTACGCCAACCTCTCGACGACCCCTTCCTCCCTCCAC
Db      532 CCCCCCTCTTACCCCTCCCTCCCTCCACACACACACCTCTT
Qy      152 CTCTCTTGACCTCTCTCTCCCAAGCCGGTTCTCTCCACCTGGCCAT
Db      592 CCCTCCCTCTCCCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC
Qy      212 CGACTCCCTCTCGCTGACCTTCTCCATCTCCACCCCTCTCCAC
Db      652 NCNCCTCCCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC
Qy      272 TAGCTCCCTCTCGACCCCTCTCTCTCCACACCTCTACCCACCGSGCTCT
Db      712 TNCCNCNNCTCTCCCTCCNCCNNCTTCCCTCCCTTCCCTACCC
Qy      332 GCTCCCTCTGGCCGCGCCGCCCCCACCCAGCAGGGCACC 371
Db      772 TCCCCCCCCNNCCNCNTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 811

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 14, 2004, 14:05:59 ; Search time 2717 Seconds
Scoring table: IDENTITY NUC

Title: US-09-669-817A-4
Perfect score: 523
Sequence: 1 atctgacccaggatcacca.....gaccgaaatggaaatggaa 523
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST *

- 1: em_estba:*
- 2: em_estchum:*
- 3: em_estrin:*
- 4: em_estmu:*
- 5: em_estcov:*
- 6: em_estapl:*
- 7: em_estrio:*
- 8: em_estrc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pin:*
- 20: em_gss_vrt:*
- 21: em_gss_fut:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rnd:*
- 26: em_gss_pig:*
- 27: em_gss_vir:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Searched: 27513289 seqs, 14931090276 residues
Searched: 1 acatgtacccaggatcacca.....gaccgaaatggaaatggaa 523
Searched: Gapext 1.0

Score: 14931090276 residues
Sequence: 1 acatgtacccaggatcacca.....gaccgaaatggaaatggaa 523
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST *

- 1: em_estba:*
- 2: em_estchum:*
- 3: em_estrin:*
- 4: em_estmu:*
- 5: em_estcov:*
- 6: em_estapl:*
- 7: em_estrio:*
- 8: em_estrc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pin:*
- 20: em_gss_vrt:*
- 21: em_gss_fut:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rnd:*
- 26: em_gss_pig:*
- 27: em_gss_vir:*
- 28: gb_gss1:*
- 29: gb_gss2:*

ALIGNMENTS

RESULT 1
AUT173524
LOCUS AUT173524 461 bp mRNA linear EST 03-APR-2002
DEFINITION AUT173524 Rice root *Oryza sativa* (japonica cultivar-group) cDNA
CLONE Clone R3620, mRNA sequence.
ACCESSION AUT173524
VERSION AUT173524.1 GR:12623311
KEYWORDS EST:
SOURCE *Oryza sativa* (japonica cultivar-group)
ORGANISM *Oryza sativa* (japonica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoidae; Oryzeae; *Oryza*.
REFERENCE 1 (bases 1 to 461)
AUTHORS Sasaki, T. and Yamamoto, K.
TITLE Rice cDNA from root (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.afrc.go.jp, URL: http://rgp.dna.afrc.go.jp/
PROJECT RGP
FEATURES location/pqualifiers
SOURCE 1. .461

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	%	Length	DB	ID	Description
No.	Score					
1	454.6	86.9	461	9	AUT173524	AUT173524 AUT173524
2	205.4	39.3	527	10	BF421014	BF421014 FMI_5_B11
3	203.8	39.0	601	14	CD463955	CD463955 ETH1_47_D
4	192.6	36.8	591	14	CA729352	CA729352 wd11C.pko

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="DNA"
/cultivar="Nipponbare"
/db_xref="Taxon:39947"
/clone="R3620"
/clone_id="rice root"
/note="Prepared from seedling root."

SEQ PRIMER: JEN REV
 HIGH QUALITY SEQUENCE STOP: 480
 POLYA=NO.
 LOCATION/QUALIFIERS
 1. .527
 /ORGANISM="Sorghum propinquum"
 /MOL_TYPE="mRNA"
 /DB_XREF="Taxon:132711"

	Best Local Similarity	Pred. No.	Indels	Gaps
Matches	457;	Conservative	0;	Mismatches 4;
Qy	10	GCAGATCACCAGAAGTCTCTCACACTCCACACTCCAGGGCACTGAGAGAGCTTCAC	69	
Db	1	GCAGATCACCAGAAGTCTCTCACACTCCACACTCCAGGGCACTGAGAGAGCTTCAC	60	
Qy	70	CGCCGCCGACCATGAGTCTCACACCGTACGCCACCTCCAGAACCCCTCT	129	
Db	61	CGCCGCCGACCATGAGTCTCACACCGTACGCCACCTCTGAGAACCCCTCT	120	

/note=Organ: floral-induced meristems; Vector: pBluescript II from Lambda ZAP II; Site 1: XbaI; Site 2: EcoRI; mature plants were placed in a growth chamber for 15 days with 16 hr darkness and 8 hr light (flowering is induced by short-day conditions); 16 days after being returned to the greenhouse under natural long days during late April/early May, meristems were harvested. The library was made from polyA RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."

Db	121	TCCCTTCCCACCGCCGTCCTCCCTGccCTTCTCCTCCACCGCCGCTTCT	180
Qy	190	CCACTTGGCCATTCTTGCCCTGACTCCCCCTCGCTGCTGACCTTTCCATCTCCACC	249
Db	181	CCACCTGCCATTCTTGCCCTGACTCCCCCTCGCTGCTGACCTTTCCATCTCCACC	240
Qy	250	CCTTCCTCCCCACGCGGCACTAGCTCTCTCTGACCCCTCCACACCCCA	309
Db	241	CCTTCCTCCCCACGCGGCCACTAGCTCTCTCTGACCCCTCCACACCCCA	300
Qy	310	CCGACCCGGCTCCAACTCGASCTGCCTCCGGCGGCGCCCAACCCAGGCCA	369
Db	301	CCGACCCGGCTCCAACTCGAGCTGCCTCCGGCGGCGCCCAACCCAGGCCA	360
Qy	370	CCAGCGGAGCTGCGCTATGTGACCGAGTCCACGGCGGAATGTAAGTGACACCG	429
Db	361	CCAGCGGAGCTGCGCTATGTGACCGAGTCCACGGCGGAATGTAAGTGACACCG	420
Qy	430	AGGACAGGCTGCGCGCGAGGAGGCTGCTCAAGTGGAG	470
Db	421	AGGACAGGCTGCGCGCGAGGAGGCTGCTCAAGTGGAG	461

RESULT 2

Qy 366 -CCGACAGGAAAGTGCACCTATGTCAACGGAGTCACGGCCGCGATGGTCAAGTGGAC 424

DEFINITION Bx2101a 5' Bill.bl A003 Floral 527 bp mRNA linear EST 28-NOV-2004
LOCUS
ACCESSION Bx2101a propinquum cDNA, mRNA sequence.
VERSION Bx2101a
KEYWORD EST
ORGANISM Sorghum propinquum
COMMENT Embryophyte, monocotyledonous plant, grass family, Sorghum
FEATURES
EXON
TRANSLATED
PROTEIN
PEPTIDE
GENE
STRUCTURE
EXON
TRANSLATED
PROTEIN
PEPTIDE
GENE
STRUCTURE

REFERENCE
AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Sudman, M., Marsala, C. and Pratt, L.H.
TITLE An EST database from Sorghum: floral-induced meristems
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpatt@uga.edu

Query Match 36.6%; Score 191.6; DB 14; Length 518;
 Best Local Similarity 68.8%; Pred. No. 9.6e-18;
 Matches 317; Conservative 0; Mismatches 120; Indels 24; Gaps 3;

ORIGIN

RESULT 5

CH231146	CA231146	518	bp	mRNA	linear	BST: 25-SEP-2003
LOCUS	SCQSF1301G02.9	Saccharum officinarum				
DEFINITION	CDNA clone SCQSF1301G02 5', mRNA sequence.					
ACCESSION	CA231146					
VERSION	CA231146.1					
KEYWORDS	EST.					
SOURCE	Saccharum officinarum					
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.					
REFERENCE	Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.					
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)					
COMMENT	The libraries that made SUCEST					
AUTHORS	Contact: Arruda, P.					
TITLE	Centro de Biologia Molecular e Engenharia Genetica					
FEATURES	1 (bases 1 to 518)					
source	Universidade Estadual de Campinas					
FEATURERS	Caixa Postal 6010, 13083-970, Campinas SP, Brazil					
source	Tel: 55 19 3788 1137					
FAX:	55 19 3788 1089					
Email:	parruda@unicamp.br					
Clone distribution:	clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bccenter.fcav.unesp.br					
Plate:	031 row: G column: 02					
Seq primer:	T7 Promoter Primer.					
Location/Qualifiers						
1. .518						
/organism="Saccharum officinarum"						
/mol_type="mRNA"						
/db_xref="taxon:4547"						
/clone="SCQSF1301G02"						
/lab_host="DB10"						
/clone_1ib="Saccharum officinarum FL3"						
/note="Organ: Base of developing inflorescence (5cm-long); Vector: pSpore1; Site 1: Sali; Site 2: NotI; An unidirectional cDNA library generated from [Base of developing inflorescence (5cm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.iad.ic.unicamp.br/public"						
ORIGIN						
CC345900	CC345900	843	bp	DNA	linear	GSS 16-MAY-2003
LOCUS	OGPAB187 ZM 0.7-1.5 KB Zea mays genomic clone ZMMaMa0350M17,					
DEFINITION	genomic survey sequence.					
ACCESSION	CC345900					
VERSION	CC345900.1					
KEYWORDS	GSS.					
SOURCE	Zea mays					
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE	1 (bases 1 to 843)					
AUTHORS	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Ronning, R., Citterk, R.W., Nusberg, A., Robbins, D. and Lakey, N.					
JOURNAL	Consortium for Maize Genomics Unpublished (2002)					
COMMENT	Contact: Cathy Whitelaw					
TIGR	9712 Medical Center Drive, Rockville, MD 20850, USA					
FEATURERS	Tel: 301-838-5843					
source	Fax: 301-838-0208					
Email:	whitelaw@tigr.org					
Seq primer:	T7					
Class:	sheared ends.					
Location/Qualifiers						
1. .843						
/organism="Zea mays"						
/mol_type="genomic DNA"						
/strain="B73"						
/db_xref="taxon:4577"						
/clone_1ib="ZMMaMa0350M17"						
/note="vector: pCpSK- Site 1: HinclII; 0.7-1.5 KB methylation filtered genomic DNA library"						

Best Local Similarity 67.1%; **Pred.** No. 1.2e-17; **Matches** 306; **Conservative** 0; **Mismatches** 138; **Indels** 12; **Gaps** 2;

/note="Organ: tassels; Vector: HybrizAP; Site: 1: EcoRI; Site 2: XbaI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybrizAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

Qy 80 CACCATAGTTCACCACTGTTACGCCAACCTCCCTCGACGACCCCTCTCCCTCCC 139
Db 222 CCCATGGCTTCCATGCCACTTCAGCTTCTTGAAAGACCCCTTC 281

Qy 140 ACGCCGCTCTCTCCCTGCCCCCTCCCTGGCTTACCCCTTCC 199

Db 282 ACCGCCGACCTTCCTCTCTCCCTGACGAGCCCTTC 341

Qy 200 ATCTTCGCTCTGACTCCCTTGCTGTGCTTACCCCTTC 259

Db 342 GTCGCCGATCCGGCTCCGGCTCTCCATTCGGGAGCTAGATGTCCTCC 401

Qy 260 CAGGCCGACCATGACTCTCCTCTCCCTGACGAGCCCTTC 310

Db 402 GCGGGCTGACCCCTCTTGCTTACCCCTGACGAGCCCTTC 461

Qy 311 CGACCGCTCTCAACTGACTGCTGACCTCCACCCGGCCACACCCAG--GCC 367

Db 462 CGACCGCTGCGACTGAGCTGCGCCACCCGGCCACCCGGCCCC 521

Qy 368 CACCGAGGAGTGCACCTAGTCAAGGAGCCACGGCCATGTCAGTGGACAC 427

Db 522 CGCCACCGGAAGGCACTATTCCTACTGAGGCGGGCCAAAGTCAGTGGATTC 581

Qy 428 CGAGGACAGCTCGCGCGGGAGAGGTGTCAGTGGGGCTGACTGACTCCC 487

Db 582 CGTGAGAGGCCGCGGCCGAGCGAGCTCAAGTGGGGGAGATCC 641

Qy 488 CTAGACGATGATCTGACCGAGTGGAAAGTGGGA 523

Db 642 CGACGACGAGCTTCGACCGAGTGGGA 677

RESULT 7

BUD92717

DEFINITION BU092717 mRNA 564 bp linear EST 29-AUG-2002.

LOCUS 946157A06_Y1_946 - tassel primordium prepared by Schmidt lab Zea

REFERENCE mays cDNA, mRNA sequence.

VERSION BUD92717.1 GI:22542279

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyt; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 564)

AUTHORS Walbot, V.

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

UNIVERSITY University Unpublished (1999)

COMMENT Contact: Walbot, V.

DEPARTMENT Department of Biological Sciences

STANFORD UNIVERSITY Stanford University

ADDRESS 855 California Ave, Palo Alto, CA 94304, USA

PHONE Tel: 650 723 2227

FAX Fax: 650 725 8221

EMAIL Email: walbot@stanford.edu

PLATE Plate: 946157 Row: A column: 06.

FEATURES source

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="taxon:4577"

/tissue_type="tassels"

/dev_stage="Just after the transition from vegetative to

inflorescence development"

/lab_host="XLOLR"

/clone_id="946 - tassel primordium prepared by Schmidt

lab"
/note="Organ: tassels; Vector: HybrizAP; Site: 1: EcoRI;
Site 2: XbaI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybrizAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."

ORIGIN

Query Match 36.3%; **Score** 189.6; **DB** 13; **Length** 564; **Matches** 305; **Conservative** 0; **Mismatches** 139; **Indels** 12; **Gaps** 2;

Qy 80 CACCATAGTTCACCACTGTTACGCCAACCTCCCTCGACGACCCCTTC 139

Db 17 CGCATGGTTCCACGCACTCTCTAGCTCGCTCAGAGCCCTTC 76

Qy 140 ACCCGCTCTCCCTCTCTGCGCTTCCTCTCCACCCCTTC 199

Db 77 ACCGGCACTCTCTCTCTGCGCTCGCTCTGAGGGCTCAGATTACCC 136

Qy 200 ATCTTCGCTCTGACTCCCTTGCTCTGCTGCTGACCCCTTC 259

Db 137 GTCCCGTCTCCATCCGCTTCTGACCTTCACCCCTTC 196

Qy 260 CACCGGAGGACACTGCTCT-----CTTGACCCCTCTCCAC 310

Db 197 GCGCGCGCTGACCCCTTCTGCGCTTACCCCGCTCAATTCTCGAGACCTCAC 256

Qy 311 CGACCGCGTCTCCACTGAGCTGCTGCGCTTCGCGCCACACCCAG--GCC 367

Db 257 CAACGTGCGCCACTATGTCACCGAGTCACCCGGCAATGGTGGACAC 316

Qy 368 CACCGAGGAGGAGGACCTATGTCACCGAGTCACCCGGCAATGGTGGACAC 427

Db 317 CGCACCGGAGGGCTTCAAGTGGAGTGGACAGTGGACTTC 376

Qy 428 CGACGACAGGCTGCGAGGGCTGCTGCAAGTGGGGCTGACTCC 487

Db 377 CGTCGAGAACGGCGGGGGAGGGCTCAAGTGGAGGCCGATAGTC 436

Qy 488 CTAGACGATGATCTGACCGAGTGGGA 523

Db 437 CGACGACGAGGCTTCAAGTGGAGTGGGA 472

RESULT 8

BUD931821

DEFINITION CA831821 mRNA 604 bp linear EST 12-DEC-2002

LOCUS 117024B07_Y1_1117 - Unigene v from Maize Genome Project Zea mays

REFERENCE CGNA, mRNA sequence.

VERSION CA831821.1 GI:26559586

KEYWORDS EST.

SOURCE Zea mays

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyt; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 604)

AUTHORS Walbot, V.

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford

UNIVERSITY University Unpublished (1999)

COMMENT Contact: Walbot, V.

DEPARTMENT Department of Biological Sciences

STANFORD UNIVERSITY Stanford University

ADDRESS 855 California Ave, Palo Alto, CA 94304, USA

PHONE Tel: 650 723 2227

FAX Fax: 650 725 8221

EMAIL Email: walbot@stanford.edu

PLATE Plate: 111024 row: B column: 07.

FEATURES source

/organism="Zea mays"
/mol_type="mRNA"
/db_xref="FAXON:4577"
/clone.Lib=1117 - UniGene V from Maize Genome Project"
/note="This library represents the unique genes found in
the fifth round of EST sequencing at Stanford University
for the maize genome project. Sequences are present from
library 946. Contigs were assembled using ZmDBAssembler
and 2 representatives from each contig were selected for
the UniGene set. All singlets were also selected."

The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: umpratt@uga.edu

Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
 the Human Genome Center, University of Tokyo Institute of Medical
 Science; plant material and RNA prepared at Texas A & M University;
 sequencing done in the Laboratory for Genomics and Bioinformatics,
 University of Georgia. Sequence ends have been trimmed to exclude
 vector and regions below Phred quality 16. Three-prime sequences
 were determined by Sanger sequencing and four-prime sequences were
 determined by T3 sequencing.

Query Match 36.3%; Score 189.6; DB 14; Length 604;
 Best Local Similarity 66.9%; Pred. No. 1.9e-17;
 Matches 305; Conservative 0; Mismatches 139; Indels 12; Gaps 2;

FEATURES
source
are presented as their reverse complements
exclude polyA.
Seq primer: SGS (CTTCGCTCTAAAGCTGG)
location/qualifiers

THE BOSTONIAN

QY	Db	61	CGCCATGGCTTCCCATCGGCACTTCATACTGCAGAACCCCTCTCCCTCC	120
QY	Db	140	ACCGCGGCTCTCTCTGCCCCCTCTCCACCCGCTTCTCCACCTGCC	199
QY	Db	121	ACCGCGGCACTCTCTCTCTCTCTCTCTCTGCGGTCTCGACCGCGTCTGAGATTA	180
QY	Db	200	ATTCITCGCCCTGACTCCCCCTTGCTGACCCTTCATCCACCCCTCTCCC	259
QY	Db	181	GTCCTCGTTCCCGCCGCTCCATTCGGGACTAGTAGTATGATCCTCTGCCCCG	240
QY	Db	260	CAGCCGACCACTAGCTCT-----GTCAGCCCTTCCTCCACACCTCAC	310
QY	Db	241	GCCGGGCTGACCCCTCTGGCTTACCCCCACCCGCTCTATCTCGAGACTCAC	300
QY	Db	311	CGACCGCGCTCCAACTCGAGCTGCGCCCTGGCCACACCCAG---GCC	367
QY	Db	301	CAACGGTGTGACGCACTCGAGGTCGCGCGACACGGCGACACGGCGAC	360
QY	Db	368	CACCGCAGGAAGTCACCTATGTCACCGAGTCCACCGGGCGCATGTCAGTGAC	427
QY	Db	361	CGCCACCCGGAAGCCACCTATGTCAGGCGCCGGCGCAAGTCAGTGGACTC	420
QY	Db	428	CGAGACAAAGCTCCGGCGCGCGAGGGCTCTAGGGAGSTGAGSTGAGTCAGATCCCC	487
QY	Db	421	CGTCAGAGAGCCGCCGGGAGGGCTCTAGGGAGGGCGAGATAAGTC	480
QY	Db	488	CTAGCAGATGGATCTGACCGCAAGTGAAGTGGAA	523
QY	Db	481	CGACGACGAGGGCTTGACCGCAACTGGAAAGTGGAA	516
RESULT 9				
LOCUS	CD429988	509	bP mRNA linear EST	03-JUN-2003
DEFINITION	ETH1_16_D10_91_A002 Ethylene-treated seedlings Sorghum bicolor cDNA clone			
ACCESSION	ETH1_16_D10_A002	5', mRNA sequence.		
VERSION	CD429988.1			
KEYWORDS	EST.			
SOURCE	Sorghum bicolor (sorghum)			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.			
REFERENCE	1 (bases 1 to 59)			
AUTHORS	Corondonier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C., Sun,F., Sullivan,R., Shan,M., Salzman,R., Chua,Tun,N., Gonzalez,M., Iane,S., Miller,V., Nanda,P., Olasehede,O., Eastman,A., and Pratt,L.H.			
TITLE	An EST database from sorghum: 1-aminoacyclopropane-1-carboxylic acid (ACC)-treated seedlings			
JOURNAL	Unpublished (2003)			
COMMENT	Other ESTs: ETH1_16_D10.b1 A002 Contact: Corondonier-Pratt MM Laboratory for Genomics and Bioinformatics			

```

/mol type="mRNA"
/cultivar="BX623"
/db xref="taxon:4558"
/clone="ETH1_16_D10_A02"
/lab host="Df1B-Tl phage-resistant E. coli"
/clone lib="ethylene-treated seedlings"
/note="vector: pMB18s-FL3; Site 1: XbaI; Site 2: XbaI; The library was prepared from polyA+ RNA from seedlings grown in hydroponic culture. At 8 days of age, medium was supplemented with 0.5 mM 1-amino-cyclopropane-1-carboxylic acid (ACC) to induce endogenous ethylene (ETH) production. Roots and shoots were harvested after 27 and 72 hr and isolated. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pMB18s-FL3 vector (5'-prime DraIII site is CACATGTC, 3'-prime DraIII site is CACCATG). XbaI excises the cDNA insert."

```

ORIGIN

Query	Match	Score	DB	Length
QY	Best Local Similarity	33.0%	172.8	509
QY	Best, Local Similarity	68.0%	Pred. No. 4.3e-15	
QY	Matches	291	0	Mismatches 122; Indels 15; Gaps 3;
QY	80	CACCAAGAGTCTCACCGTTAGGCCACCTCTCCACCGGGTTCTCC	139	
QY	82	CGCCATAGCTTCCACCGCAGCTCCGCTCCACCTTGACGACCCCTCTCC	141	
QY	140	ACCGCG-----TCTCTCTCTCTGCCTCTCTCCACCGGGTTCTCC	193	
QY	142	ACCGCGCACCTCTCATCTCCACCTTGACGGGTCTAGACCC	201	
QY	194	CTGCCCATCTCGCTGACTCCCTT-----CGCTGTGACCCCTTCATCCA	247	
Db	202	GRCGCGTCCGAGACTGGAGACTCCCTCTGCGCTCGACCGGGCTCGA	261	
QY	248	CCCTCTCTCCACCGGACACTGACTCTCTCTGACCCCTCTCTCCACCC	307	
Db	262	CCCTCTCTCCACCGGACACTGACTCTCTCTGACCCCTCTCTCCACCC	321	
QY	308	CACGACCGCGCTCCAACTGGAGGTGGCTCGCGCGGCCAACCGGG--	365	
Db	322	CACGACCG	381	
QY	366	-CCACCAAGCAGGAAGTCACCTATCTACCGAGTCACCGCCCGATGTC	424	
Db	382	CCCGCCACCCGAAAGCGACCTACCTCACCGCCACGGGACGCTCAAGTGGAC	441	
QY	425	CACGAGACAGCAGCTGGCGCGCGAGGGTGTCAAGTGGAGCTGCTGACTC	484	
Db	442	CTGTGTCAGAGAACCGCGCGGGAGCGACGCTCATGGAGGCCAGATCAAGTC	501	
QY	485	CCCTTACG 492		
Db	502	CCCGAGCG 509		

Db	Matched	267;	Conservative	0;	Mismatches	114;	Indels	15;	Gaps	3;
Qy	80	CACCATGAGTTTCAACCGGTAGGCCAACCTCTCGACGACCCCTTCCTCCCTCCC	139							
Qy	263	GGCCTATGCTTCAACCGCACCTCTCGACAAACCTCTTCCTCC	322							
Db	308	CCCCGGGCGCTCTCCAACTCTGAGSTCGCTCGCCGCGCCCCGCC	379							
Qy	366	-CCCACAGAGGAATGGCACTATTCACCGAGTCACCGCCGATGTCAGGGAC	424							
Db	380	CCCCGGGACCCGGAACGACTACGGTACCGACGCGGGCGGAGGAGGAC	439							
Qy	425	CCCCGGGACAGCTGGCGGGAGGGTCTAAGTGG	468							
Db	440	CTCTGTGAGAGCCGGGGAGGAGCTAGGTGG	483							
RESULT 12										
LOCUS	CD213737	658 bp	mRNA	linear	EST	20-MAY-2003				
DEFINITION	HS1_52_G09_g1_A012	Heat-shocked seedlings	Sorghum bicolor	cDNA						
COMMENT	clone HS1_52_G09_A012 5'	, mRNA sequence.								
VERSION	CD213737									
EST	CD213737.1	GI:30952107								
AUTHORS										
REFERENCE										
ORGANISM	Sorghum bicolor (sorghum)									
TITLE										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 13										
LOCUS	CA168889	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.9	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168889									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 14										
LOCUS	CA168890	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.10	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168890									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 15										
LOCUS	CA168891	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.11	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168891									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 16										
LOCUS	CA168892	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.12	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168892									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 17										
LOCUS	CA168893	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.13	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168893									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 18										
LOCUS	CA168894	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.14	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168894									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 19										
LOCUS	CA168895	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.15	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168895									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 20										
LOCUS	CA168896	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.16	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168896									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 21										
LOCUS	CA168897	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.17	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168897									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 22										
LOCUS	CA168898	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.18	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168898									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 23										
LOCUS	CA168899	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.19	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168899									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCD									
TIME										
JOURNAL										
COMMENT										
FEATURES	source									
FEATURES										
FEATURES	source									
RESULT 24										
LOCUS	CA168900	662 bp	mRNA	linear	EST	24-SEP-2003				
DEFINITION	SCUFSB010D05.20	SR1	Saccharum officinarum	cDNA clone	SCUFSB010D05					
COMMENT	5'	mRNA Sequence.								
VERSION	CA168900									
KEYWORDS										
ORGANISM	Saccharum officinarum									
REFERENCE										
AUTORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta;									

double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 29.1%; Score 152; DB 14; Length 662;
Best Local Similarity 62.6%; Pred. No. 3. 8e-12; Mismatches 145; Indels 24; Gaps 2;
Matches 233; Conservative 0; /note="Organism: Developed inflorescence and rachis (20cm-long); Vector: pspotl; Site 1: SalI; Site 2: NotI; /lab_host="PH10B"/clone=SCRFL4008BG02"

Qy 93 CACCAACGGTTAGCCGACTCTCGAGACCCCTCTCCCTCCACGCCGTCTC 152
Db 46 CTCCACGGCGTGAAGAGCACCCCTCTCCATCTTCAGAACACC 105

Qy 153 TCTCTGCCTCTCTCCACCCACCCGGGTTCTGACCTGCCATTCTCGCCCTC 212
Db 106 TCTGCTGCCTCATCTCCACCTCCGCTCTCCGACCTCCGAC 165

Qy 213 GACTCCCTCTCGTGCTGACCCCTCTCCACCTCCGACCTCCGAC 272

Db 166 GACCAACCCCTCTACCCACCTCTCCGCGCTCGCTCACCTCCCTCCGATGAC 225

Qy 273 AGCCCTCTCTGACCCCTCTCTCCACCCCTAC 311

Db 226 ACCTTGACCTCGACCTCTCTCCGCGCCACGCCGCGCCCCGGCTTC 285

Qy 312 GACCGGGCTCCAACTGAGCGAGCTGGCCGGCCGGCCACCCCTGGCCACC 371

Db 285 GACTTGACCCCTCTCTCCGAGCGCTGGCCACCCGGTCTCGCCCG 345

Qy 372 AGCAGGAATGCACTATGCAAGGAGTCAACGGCCATGCACTGAGAAGTCAAGTGGACCGAG 431

Db 345 CGTCCGAATTAACCTAGCCCTGGAGGCCAGGGAGAAGTCAAGTGGACCGAG 405

Qy 432 GACAAGCTGGCGCGGGAGGGGTCTCAAGTGGAGGCTGACTCCCTAC 491

Db 406 GACAGGCC--CGCGGAGACCGACCTCAAGTGGAGGAGCTCAGGAGCCCTAC 462

Qy 492 GACGATGGATTGACCGCAAGTGAAGTGGGA 523

Db 463 CACGACGGCTCGACCGCAAGTGGAGTGGGA 494

FEATURES
source
Location/Qualifiers
1. . 657
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="Taxon:4547"
/clone=SCRFL4008BG02"
/note="Organism: Developed inflorescence and rachis (20cm-long). DNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Query Match 28.8%; Score 150.4; DB 14; Length 657;

Best Local Similarity 62.4%; Pred. No. 6. 4e-12; Mismatches 146; Indels 24; Gaps 2;
Matches 282; Conservative 0; /note="Organism: Developed inflorescence and rachis (20cm-long). DNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

Qy 93 CACCAACGGTTAGCCGACTCTCGAGACCCCTCTCCCTCCACGCCGTCTC 152

Db 71 CTCCACGGCTGAGCACCCCTCTCCATCTTCACCTCCCTCCACACACC 130

Qy 153 TCTCTGCCTCTCTCCACCCCTCTCCGCTCTCCGCTCTCCGAC 212

Db 131 TGCCTGCGCTACATCTCCGCTCTCCGACACCGCTTCCTCTCGAC 190

Qy 213 GACTCCCTCTGCGCTCACCCCTCTCCACCCCTCTCCACCTCAC 272

Db 191 GACCAACCCCTCTACCCACCTCCCTGGCCGCTCTGACCTCCCTCCCGATGAC 250

Qy 273 AGCTCTCTGACCCCTCTCCACCCACCTCAC 311

Db 251 ACCTTGACCCCTCTCCACCCACGGCCGGCCGGGCTCGCCCTC 310

Qy 312 GACCGGGCTCTCCACCTCCGAGCTGCCCTCTGGCCGGCCACACCCAGGGAC 371

Db 311 GACTTGACCCCTCTCCGAGGCGCTGGCCACCGGTCTGGCCCGCCGGCCCG 370

Qy 372 AGCAGGAATGCACTATGCAAGGAGTCAACGGCCATGCACTGAGAAGTCAAGTGGACCGAG 431

Db 371 GTCGAAGATGACCTACCTCTGGAGGCCACGGGAGGCTGGAGGCG 430

Qy 432 GACAAGCTGGCGCGGGAGGGGTCTCAAGTGGAGGCTGAGCTGAGCTCCCTAC 491

Db 431 GACAGGCC--CGCGGAGACCGACCTCAAGTGGAGGAGCCGAGCTGGAGGAC 487

Qy 492 GACGATGGATTGACCGCAAGTGAAGTGGGA 523

Db 488 CACGACGGCTGACCGCAAGTGGAGTGGGA 519

RESULT 14
CA221736 CA221736 657 bp mRNA linear EST 25-SEP-2003
DEFINITION SCRFL4008BG02.G FL4 Saccharum officinarum cDNA clone SCRFL4008BG02
ACCESSION CA221736
VERSION CA221736.1 GI:35276347
KEYWORDS EST
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
REFERENCE Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
AUTHORS 1 (bases 1 to 657)
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arudra P
Centro de Biologia Molecular e Engenharia Genética
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bccc.cfev.unesp.br>
Plate: 008 row: G column: 02
Seq primer: T7 Promoter Primer.

TITLE

The libraries that made SUCRST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

JOURNAL

COMMENT

Centro de Biologia Molecular e Engenharia Genética
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
<http://www.bcccenter.fcav.unesp.br>

Plate: 050 row: G column: 08

Seq primer: T7 Promoter Primer.

FEATURES location/Qualifiers

1. . 667
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCB0PL450G08"
/lab_host="BH10B"
/clone_lib="FL4"
/note="Organ: Developed inflorescence and rachis
(20cm-long); Vector: pSPori1; Site 1: Sali; Site 2: NotI;
An unidirectional DNA library generated from [Developed
inflorescence and rachis (20cm-long)]. cDNA was prepared
from polyA+ mRNA using SuperScript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
<http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Job time : 2733 secs

Query Match 28.8%; Score 150.4; DB 14; Length 667;
Best Local Similarity 62.4%; Prcd. No. 6.4e-12; Mismatches 282; Conservatve 0; Mismatches 146; Indels 24; Gaps 2;

```

QY  93 |||||GACCAACCGTTAACGCCACCTCTGTGACGACCCCTCTTCCACCGCGTCC 152
    |||||CTCCACACGGCTGAGACGACGCGCTTCATCCTCTTCCACCGACACCAC 141
Db  82 |||||GACCAACCGTTAACGCCACCTCTGTGACGACCCCTCTTCCACCGCGTCC 212
    |||||TCTCTCTGACCTCTCTCCACCGACGCGCTTCATCCTCTTCCACCGAC 201
QY  153 |||||TCTCTCTGACCTCTCTCCACCGACGCGCTTCATCCTCTTCCACCGAC 212
    |||||TGTCTTGACCTACTCTCCACCGACGCGCTTCATCCTCTTCCACCGAC 201
Db  142 |||||TGTCTTGACCTACTCTCCACCGACGCGCTTCATCCTCTTCCACCGAC 201
    |||||TGTCTTGACCTACTCTCCACCGACGCGCTTCATCCTCTTCCACCGAC 201
QY  213 |||||GACTCCCTCTTGTGCTGCTGACCCCTTCACCATCCTCCACCGCTTCC 272
    |||||GACTCCCTCTTGTGCTGCTGACCCCTTCACCATCCTCCACCGCTTCC 261
Db  202 |||||GACCAACCCCTCTTGTGCTGCTGCTGCTGCTTCCACCGCTTCCCGATGAC 311
    |||||GACCAACCCCTCTTGTGCTGCTGCTGCTGCTTCCACCGCTTCCCGATGAC 311
QY  273 |||||AGCTCCCTCTTGTGCTGCTGACCCCTTCACCATCCTCCACCGCTTCC 321
    |||||AGCTCCCTCTTGTGCTGCTGACCCCTTCACCATCCTCCACCGCTTCC 321
Db  262 |||||AGCTCCCTCTTGTGCTGCTGACCCCTTCACCATCCTCCACCGCTTCC 321
    |||||AGCTCCCTCTTGTGCTGCTGACCCCTTCACCATCCTCCACCGCTTCC 321
QY  312 |||||GACCGCGTCTTCACTCGAGGTCCTCGACCCCGCCACACCGCGCC 371
    |||||GACCGCGTCTTCACTCGAGGTCCTCGACCCCGCCACACCGCGCC 371
Db  322 |||||GACTTGACCCCTTCTCTGACCGCGCTGCGACCGCCACCGCTTCC 381
    |||||GACTTGACCCCTTCTCTGACCGCGCTGCGACCGCCACCGCTTCC 381
QY  372 |||||AGCAGGAGTGTACCTATGTCACCGCAGTCACCGCGCGATGGTCAAGTGG 431
    |||||AGCAGGAGTGTACCTATGTCACCGCAGTCACCGCGCGATGGTCAAGTGG 431
Db  382 |||||CGTCGCAAGTACCTACCTACCGCCAGGGCGACGGCAAGGGCAAGGG 441
    |||||CGTCGCAAGTACCTACCTACCGCCAGGGCGACGGCAAGGGCAAGGG 441
QY  432 |||||GACAACGCTCGCGCGCGAGGGTGTCAAGTGGAGGTGAGTCACCTCC 491
    |||||GACAACGCTCGCGCGCGAGGGTGTCAAGTGGAGGTGAGTCACCTCC 491
Db  442 |||||GACGATGATTGACCGCAAGTGGAGTGGGA 523
    |||||GACGATGATTGACCGCAAGTGGAGTGGGA 523
Db  492 |||||GACGAGGGCTTGGACCGCAAGTGGAGTGGGA 530
    |||||GACGAGGGCTTGGACCGCAAGTGGAGTGGGA 530

```